A NEW METHODOLOGY FOR ANALYZING AND PREDICTING U.S. LIQUEFIED NATURAL GAS IMPORTS USING NEURAL NETWORKS

A Thesis

by

MATTHEW SCOTT BOLEN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2005

Major Subject: Petroleum Engineering

A NEW METHODOLOGY FOR ANALYZING AND PREDICTING U.S. LIQUEFIED NATURAL GAS IMPORTS USING NEURAL NETWORKS

A Thesis

by

MATTHEW SCOTT BOLEN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	Richard Startzman
Committee Members,	Maria Barrufet
	Detlef Hallermann
Head of Department,	Steve Holditch

August 2005

Major Subject: Petroleum Engineering

ABSTRACT

A New Methodology for Analyzing and Predicting

Liquefied Natural Gas Imports Using Neural Networks.

(August 2005)

Matthew Scott Bolen, B.S., Virginia Polytechnic and State University

Chair of Advisory Committee: Dr. Richard A. Startzman

Liquefied Natural Gas (LNG) is becoming an increasing factor in the U.S. natural gas market. For 30 years LNG imports into the U.S. have remained fairly flat. There are currently 18 permit applications being filed in the U.S. and another 10 permit applications being filed in Canada and Mexico for LNG import terminals. The EIA (Energy Information Agency) estimates by 2025 that LNG will make up 21% of the total U.S. Natural Gas Supply.

This study developed a neural network approach to forecast LNG imports into the U.S. Various input variables were gathered, organized into groups based on similarity, and then a correlation matrix was generated to screen out redundant variables. Since a limited number of data points were available I used a restricted number of input variables. Based on this restriction, I grouped the input variables into four different scenarios and then generated a forecast for each scenario. These four different scenarios were the $/MMBTU model, natural gas energy consumption model, natural gas consumption model and the energy stack model.

The standard neural network approach was also used to screen the input variables. First, a correlation matrix determined which variables had a high correlation with the

output, U.S. LNG imports. The ten most correlated input variables were then put into correlation matrix to determine if there were any redundant variables. Due to the lack of data points only the five most highly correlated input variables were used in the neural network simulation.

A number of interesting results were obtained from this study. The energy stack model and the consumption of natural gas forecasted a non-linear trend in U.S. LNG imports, compared to the linear trend forecasted by the EIA. The energy stack model and consumption of natural gas model predicted that in 2025 U.S. LNG imports will be about 6.5 TCF, while the other three models prediction is about three times as less. The energy stack model is the most realistic model due its non-linear trend, when the rapid increase of LNG imports is going to occur, and the quantity of U.S. LNG imports predicted in 2025.

DEDICATION

This thesis is dedicated to my parents.

Without them I wouldn’t be where I am today.

ACKNOWLEDGEMENTS

I would like to thank Dr. Richard A. Startzman for all his support during my coursework here at Texas A&M. I would like to thank Maria A. Barrufet and Detlef Hallerman for serving as committee members. I would like to thank all the students and faculty at the Harold Vance Department of Petroleum Engineering for an enjoyable experience.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

1.1 Motivation

Natural gas is quickly becoming one of the most integrated fuels in our society. From chemical feed stocks to electrical generation, natural gas is considered one of the most important fuels today. The EIA (Energy Information Agency) estimates that by 2025 the U.S. consumption of natural gas will reach 30.7 TCF[1]. In order to meet this need, several events must take place. The first is drilling more wells. The more wells we drill the more gas we can recover. Secondly, we can cut back on our use of natural gas through conservation. Finally, we can import more natural gas to meet our needs.

I believe the easiest and most economical way to satisfy our need is to import more natural gas. We are already importing 3.2 TCF from Canada and importing 0.51 TCF by LNG[2]. Canada in the future will not be able to export enough natural gas to meet the U.S. demand due to declining production and increasing demand[1]. Therefore, LNG becomes an important source of meeting our natural gas needs.

Determining how much LNG we will import to satisfy this need is an important question. The EIA uses its National Energy Modeling System, NEMS, to forecast LNG imports3. NEMS determines the supply, imports, and demand for all energy sources.

This thesis follows the format and style of the *Journal of Petroleum Technology*.

The Natural Gas Transmission and Distribution Module, NGTDM, is a module under NEMS which uses the supply, imports, and demand data as inputs to determine natural gas end-use and wellhead prices, and determines the flow patterns of natural gas through the regional interstate network. LNG imports are determined in a series of steps.

1. A least cost transportation algorithm is run multiple times to establish LNG supply curves at the beginning of each NEMS forecast year.
2. The least cost transportation algorithm establishes six points on the LNG supply curve for each of the twelve regions.
3. Market prices from the previous iteration, which were determined by the NGDTM, are used to evaluate the supply curves.
4. The values determined for LNG imports, from each supply curve, are totaled to give the amount of LNG imported for that year.

1.2 Research Objectives

The objectives of this thesis are as follows:

1. Develop a methodology using four neural networks models to predict LNG imports into the U.S.
2. Develop a comparison forecast based on current LNG terminal capacity.
3. Compare four neural network scenarios, and the LNG terminal capacity model to the EIA forecast.

1.3 Organization of This Thesis

This thesis is organized into a total of seven chapters, including this introduction.

Chapter II will focus on LNG. Topics in this chapter will include a brief overview of the natural gas industry, different aspects of LNG, and some issues concerning future supply of natural gas.

Chapter III will focus on the background of neural networks. This chapter will cover advantages and disadvantages of neural networks, current uses of neural networks and features of the neural network used in this study.

Chapter IV will focus on the development of the LNG data including data normalization and correlation that will be used in neural network model described in Chapter V.

Chapter VI describes the neural network model results and compares results against other published data. Chapter VII provides the summary, conclusions and recommendations.

CHAPTER II

LNG

2.1 U.S. Natural Gas Industry

Historically, natural gas was simply a low profit by-product of exploration, often flared or vented to the atmosphere. Today, natural gas has become an integral part of the infrastructure that exists in the U.S. The natural gas industry encompasses a wide range of participants. E&P companies produce natural gas and send it through pipelines. Transmission companies move natural gas from domestic and foreign sources to where it is needed. Natural gas is then used by residential consumers for heat and cooking and industrial consumers for feedstock.

Fig. 2.1 shows the distribution of natural gas production in the U.S. From the graph we see that unconventional production has become the largest supply of natural gas in the U.S. This is due to the increase interest in coal bed methane, shale and tight gas[1]. However, conventional production is declining and is expected to continue its decline into the future.

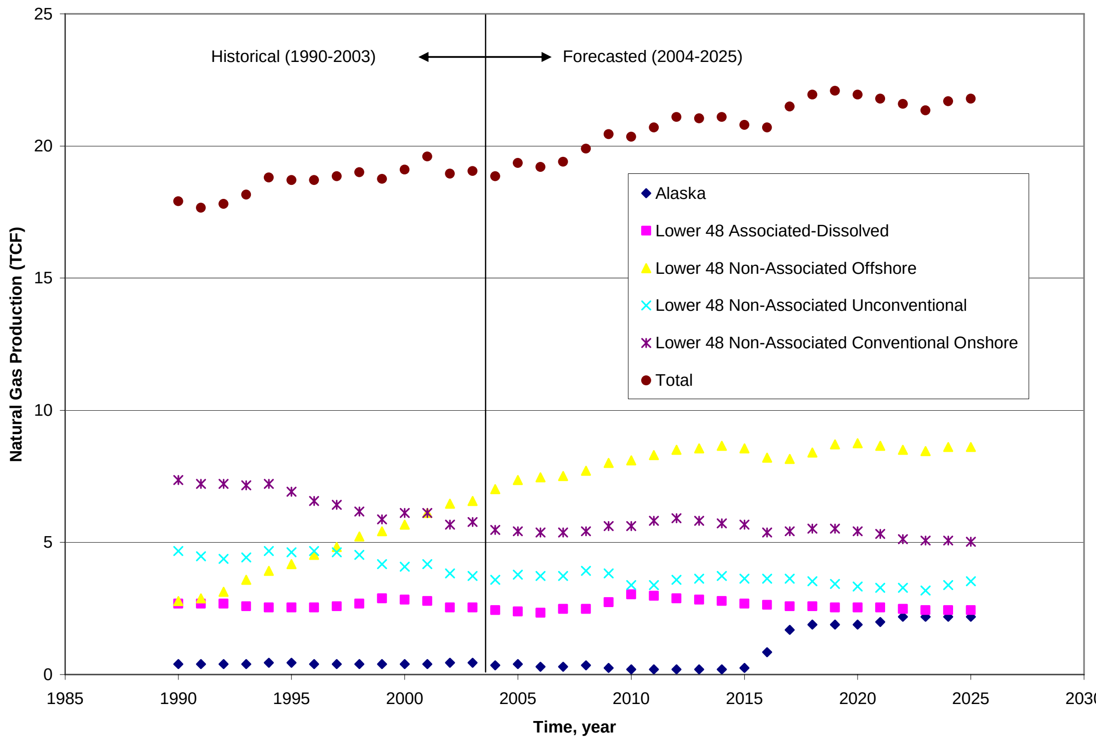

Figure 2.1 Historical Natural Gas Production with Forecast Through 2025 (EIA, AEO[1] 2005)

Fig. 2.2 shows the pipeline network for natural gas based on capacity. The figure shows that a significant portion of natural gas is transported from Henry Hub in LA to the Northeast. Also a significant portion of natural gas is supplied by Canada to the West Coast, the Midwest and the Northeast. Pipelines are vital in that they supply areas of the U.S. where natural gas is not readily abundant.

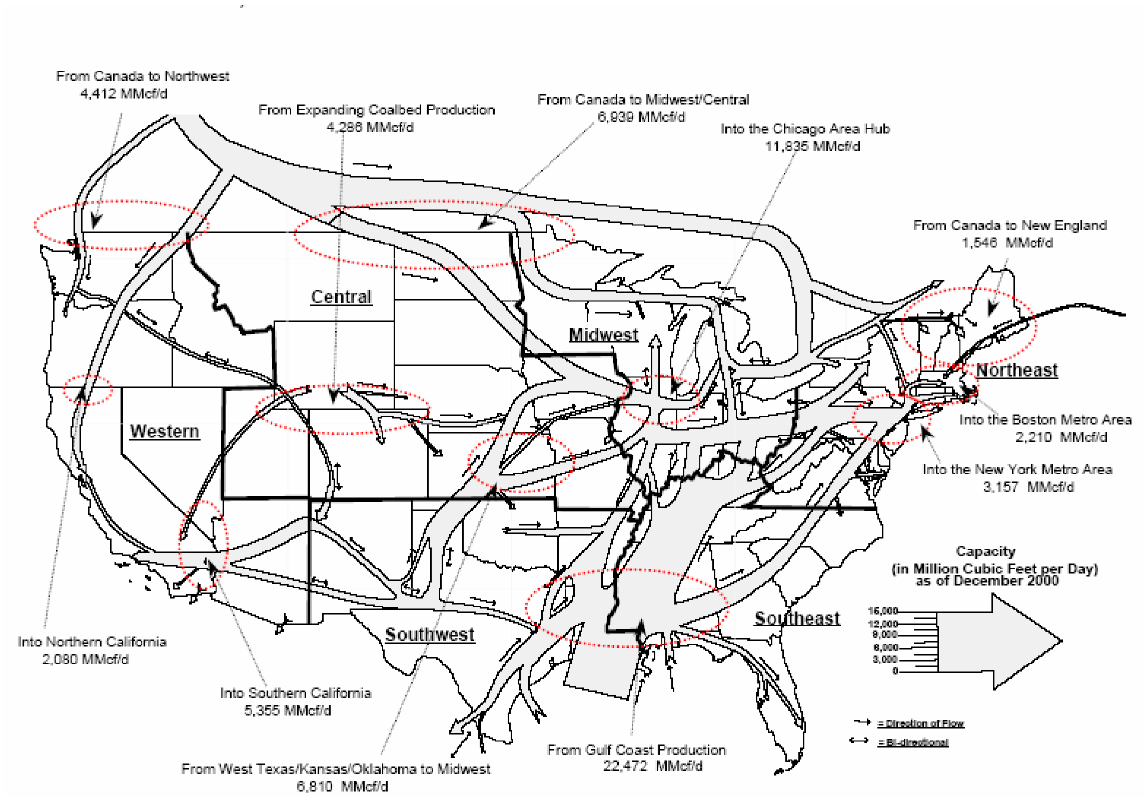

Figure 2.2 Major Natural Gas Pipeline Transportation Routes and Capacity Levels (EIA[4], 2001)

There are several consumption sectors that rely on natural gas on a daily basis. **Fig 2.3** shows the historic and forecasted consumption of natural gas for these different sectors. Industrial consumption of natural gas has declined in the past few years, but is expected to climb due to a growing use by certain industrial sectors. Iron, steel, and aluminum industries are expected to decrease consumption of natural gas while metal-based durables, petroleum refining, bulk chemicals, and food industries are expected to increase consumption[1]. Residential and commercial consumption of natural gas is expected to steadily increase due to the rise in demand from each sector. Transportation consumption of natural gas is expected to remain flat over the long term. This is due to the rise in popularity and convenience hybrid vehicles. Electrical generators consumption of natural gas is expected to rise dramatically over the next few years.

There are several reasons why this is expected to take place. These reasons compared with coal electrical generation include: lower capital costs, higher fuel efficiency, shorter construction lead times, and lower emissions[1].

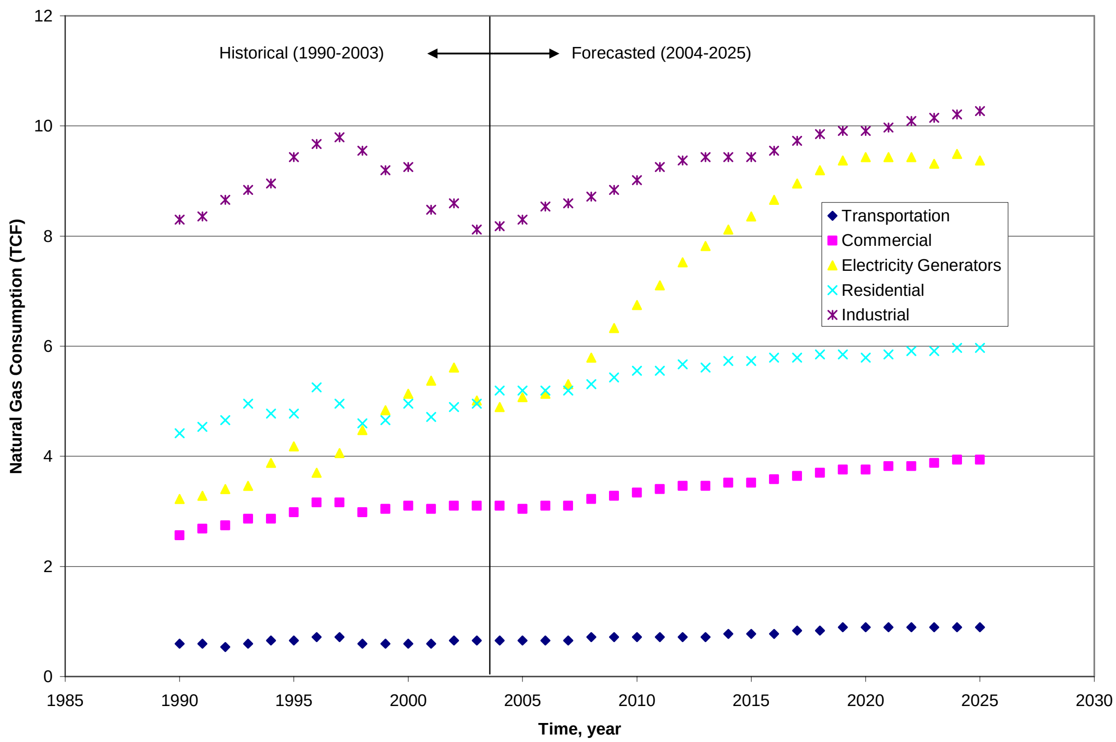

Figure 2.3 Historical Natural Gas Consumption by Sector with Forecast Through 2025 (EIA, AEO[1] 2005)

The future of the energy supply for the U.S. remains a big concern, so much so that President George W. Bush has called upon Congress to pass an energy bill. As **Fig. 2.4** shows there is an increasing gap between energy production and energy consumption in the U.S. The only way to make up that difference is by importing fossil fuels to meet the growing demand.

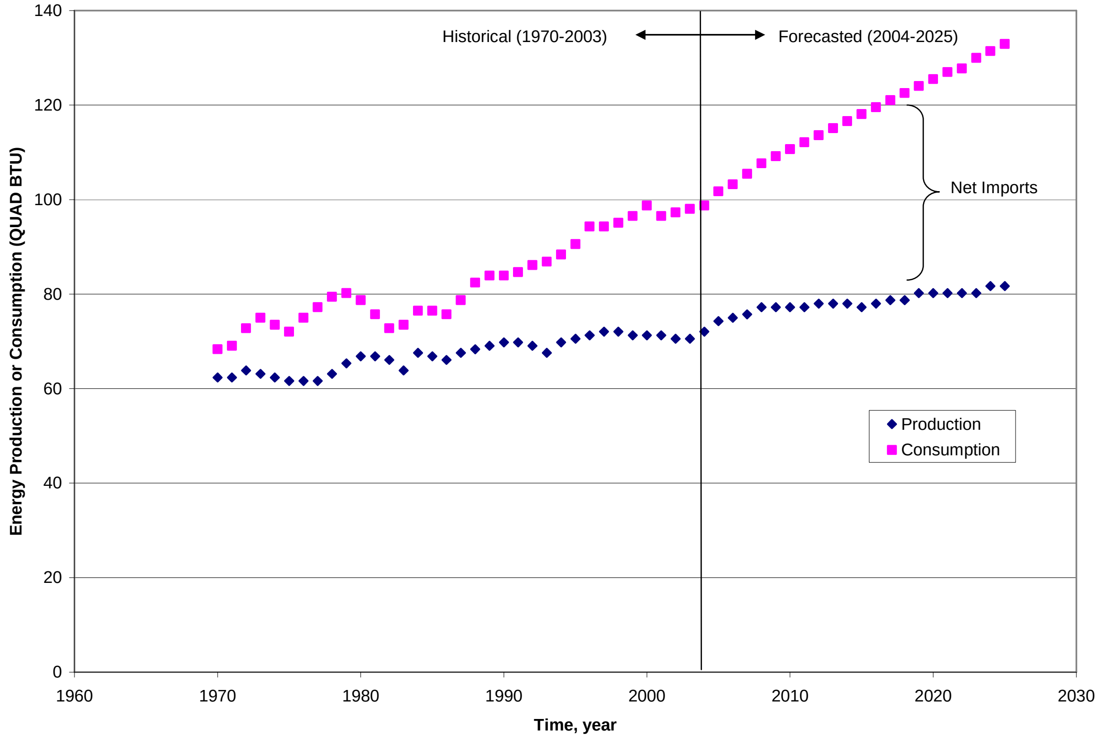

Figure 2.4 Historical Total Energy Production and Consumption with Forecast Through 2025 (EIA, AEO[1] 2005)

Fig 2.5 shows the historical and forecasted imports of natural gas into the U.S. by source. Several important trends can be drawn from this figure. First, looking at Mexico we that that natural gas imports have remained fairly flat and in the future Mexico is predicted to become a net importer of natural gas from the U.S. This trend is due to the fact that the national oil company of Mexico, PEMEX, had about 60 percent of its revenue going to the Mexican government in 2003[5]. This trend is expected to continue into the future. With sixty percent of its $56.3 billion revenue going to the Mexican government, PEMEX is unable to meet its country's demand for energy and therefore will continue to be a net importer of energy, especially natural gas.

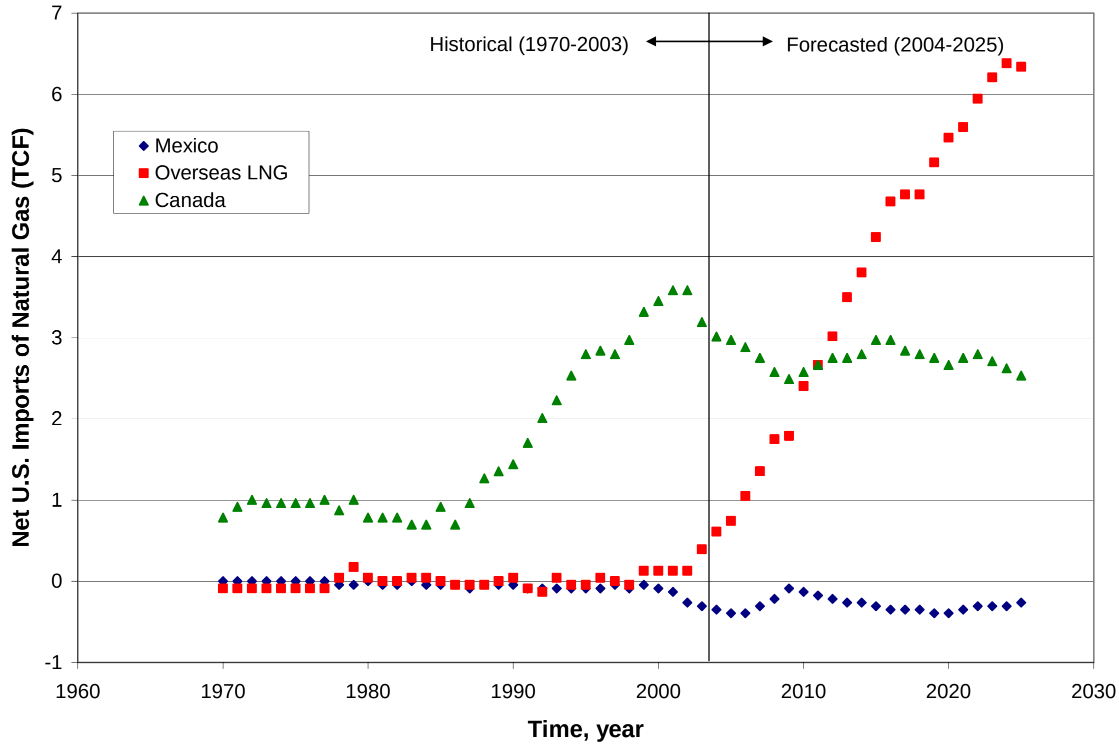

Figure 2.5 Historical Net U.S. Imports of Natural Gas with Forecast Through 2025 (EIA, AEO[1] 2005)

Looking at natural gas imports from Canada we see that historically they have been increasing. However, over the past few years imports from Canada have been decreasing and are forecasted to continue this trend. This decrease is due to the fact that the consumption of natural gas is expected to increase more rapidly than production[1].

Looking at LNG imports we see that LNG imports have remained flat for quite a while. Recently, LNG imports have increased and are forecasted to continue this trend. This is most evident by the number of LNG terminal permit applications that have been filed. According the EIA, there have been 19 permits filed to build U.S. LNG terminals, with one terminal permit being rejected by the city of Fall River, MA (Reference 5). Also there have been 8 permits filed in sites surrounding the U.S., which include:

Canada, Mexico, and the Bahamas. There are several reasons why LNG imports are expected to rise dramatically in the coming years. Looking back at **Fig. 2.4**, total energy consumption is expected to outpace total energy production. One way to make up that gap is to import more oil and natural gas. Also the high price of natural gas has made LNG an economical and an attractive option for importing natural gas. **Fig. 2.6** shows the cost of transporting natural gas and oil by various means. This figure shows that a distance of greater than 5000 kilometers LNG costs about $3.0/MMBTU. In 2003 the U.S. had an average price of $6.86/MMBTU for natural gas[2]. This difference in price shows that LNG is a very economical means for supplying natural gas to the U.S.

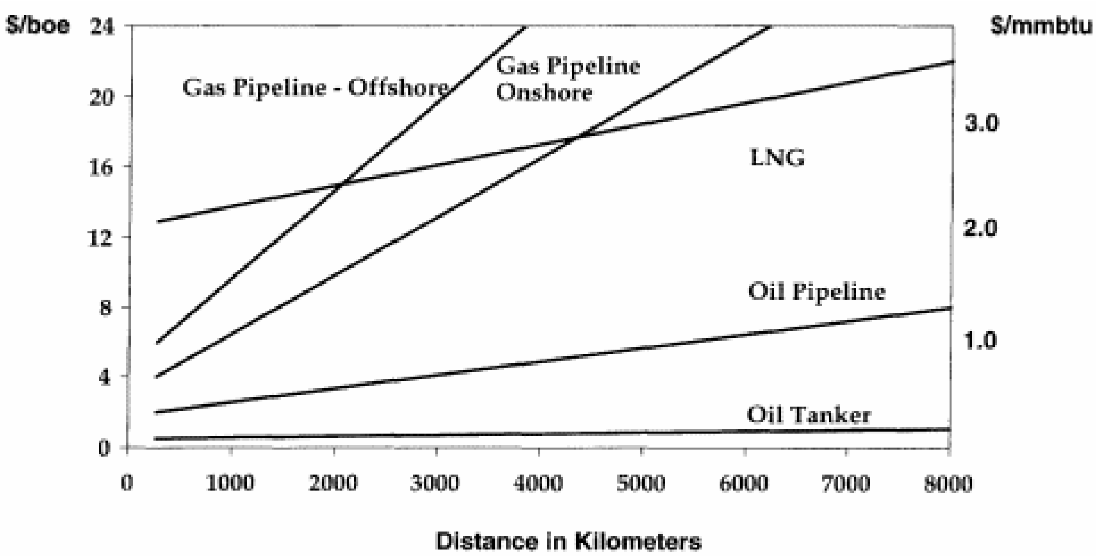

Figure 2.6 Oil and Gas Transportation Costs (Greenwald[6], 1998)

2.2 Aspects of LNG

When deciding to undertake a LNG project many factors have to be taken into account. In this section I will give a general background of LNG, and then discuss the

LNG supply chain. The LNG supply chain includes the liquefaction, transportation and regasification of LNG. Next I will bring to light some of the governmental issues surrounding LNG, and finally discuss the economics involved in an LNG project.

Natural gas is cooled to -260^{o}F in atmospheric conditions to form LNG. By changing natural gas into a liquid form, this reduces the volume of the natural gas by a factor of about 610. This allows for the economical transportation of natural gas over great distances. As of late 2003, there were a total of 151 LNG tankers operating throughout the world and another 55 tankers under construction[7]. Most of the LNG activity in the world is centered around Asia with Japan receiving about 48% of the world's LNG imports in 2002[7].

The LNG chain represents the flow of natural gas from the wellhead to the ships to the consumers in the distant country. **Fig. 2.7** shows the LNG chain and the roles that the sellers and buyers play in the chain. From this figure we can see that natural gas is produced, then liquefied, and sent on LNG tankers to the specific country. The seller or buyer can have control over the LNG tankers, which determines the type of contract used in the agreement. If the responsibility for shipping LNG is on the seller's then the two parties will enter into an ex-ship or cost, insurance, freight (CIF) contract[7]. If the buyer is responsible for the shipping then the two parties enter into a free on board (FOB) contract.

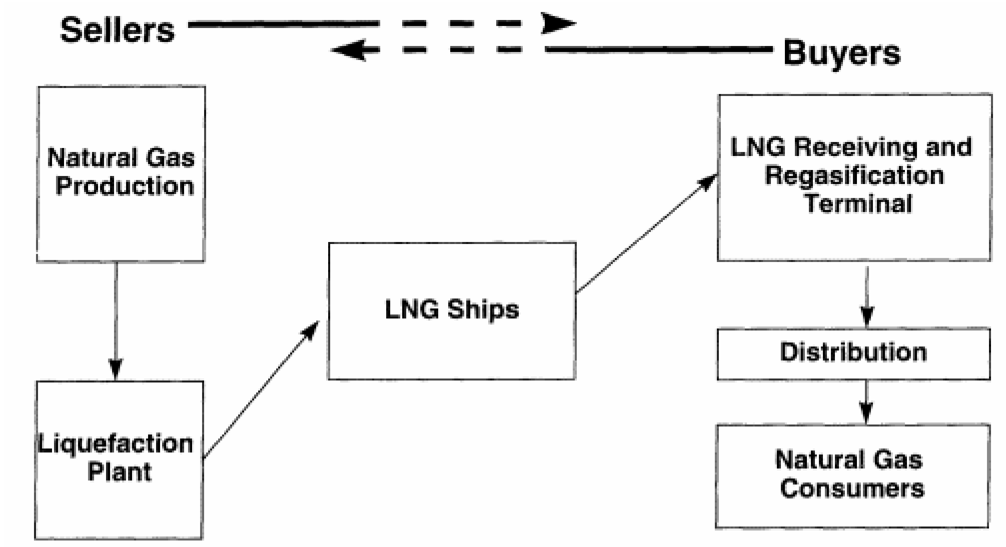

Figure 2.7 LNG Supply Chain (Greenwald[6], 1998)

The LNG is then received at the LNG regasification terminal where it is heated back into a gas. Next this gas is distributed throughout the pipeline network of the country, to the consumers of the natural gas. **Fig. 2.8** shows a typical schematic of an onshore LNG marine receiving terminal. **Fig. 2.9** displays a class of new offshore LNG marine receiving terminals. This terminal named Energy Bridge took its first LNG shipment 3/17/05. This new class of terminals is designed to tie into existing offshore pipelines, and make offloading LNG easier. Offloading is abetted due to the fact that the water is deeper and LNG ships have more room to maneuver. Onshore marine LNG terminals have to be located in water deep enough to handle LNG ships and also the LNG ships themselves have to deal with coastal traffic located around the receiving terminal.

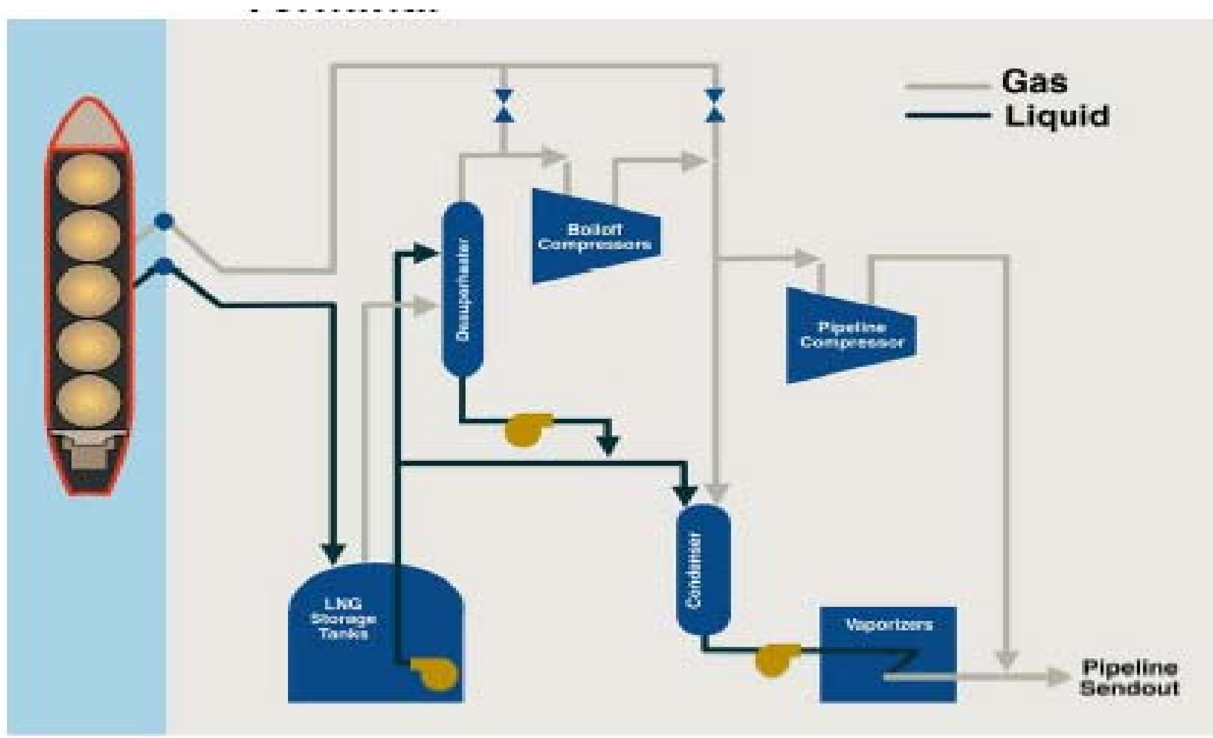

Figure 2.8 Schematic of LNG Marine Terminal (EIA[8], 2004)

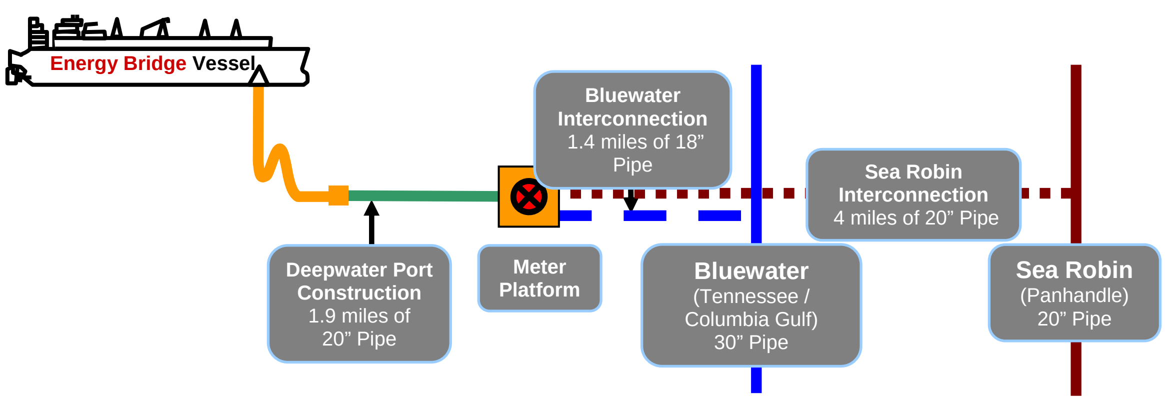

Figure 2.9 Schematic of Excelerate's Energy Bridge Vessel (Eisbrenner[9], 2004)

The Federal Energy Regulatory Commission (FERC) and the Coast Guard are the two main governmental entities that oversee the permitting for the LNG terminals. FERC regulates marine terminals built on land while the Coast Guard/MARAD (U.S. Maratime

Administration) regulates the offshore terminals. Also, onshore and offshore marine terminals are sometimes subject to local and state regulations.

In undertaking any project the economic feasibility of the project must be considered. **Fig. 2.10** shows typical cash flows for an LNG project compared to that of an oil project. An oil project hits the breakeven point twice as fast as an LNG project. However, LNG projects tend to last longer. This is due to the fact that an oil project's production profile depends mainly on the producibility of the reservoir[6]. In LNG projects, the buyer is looking for long term levels of production that can be sustained for 20 years or more[6].

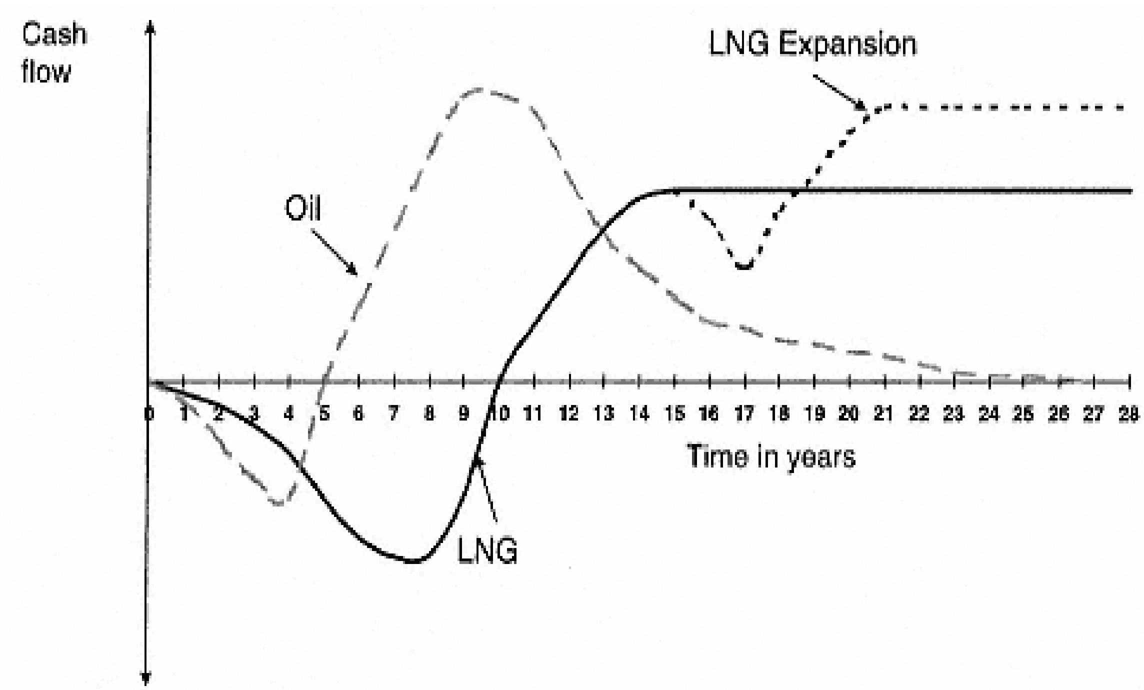

Figure 2.10 Illustrative Cash Flows for an Oil and LNG Project (Greenwald[6], 1998)

2.3 Future Natural Gas Supply

The future of natural gas supply to the U.S. will be influenced by a number of factors. These factors include:

1. Increased demand from developed and developing nations such as Japan, China, and India.
2. The proposed construction of the Alaskan pipeline from the Alaskan North Slope.
3. Regulatory hurdles inherent to onshore and offshore marine terminals
4. Competition for steel and other commodities that go into making LNG tankers.

The U.S. is currently and will continue to be in competition with Japan, China, and India for supplies of natural gas. Although the U.S. has been importing LNG for about 30 years, countries such as China and Japan have been importing LNG on a much bigger scale than the U.S. In **Fig. 2.11** we see the world wide movement of LNG in 2002. As mentioned earlier, most of the LNG activity is in the Asia/Pacific region.

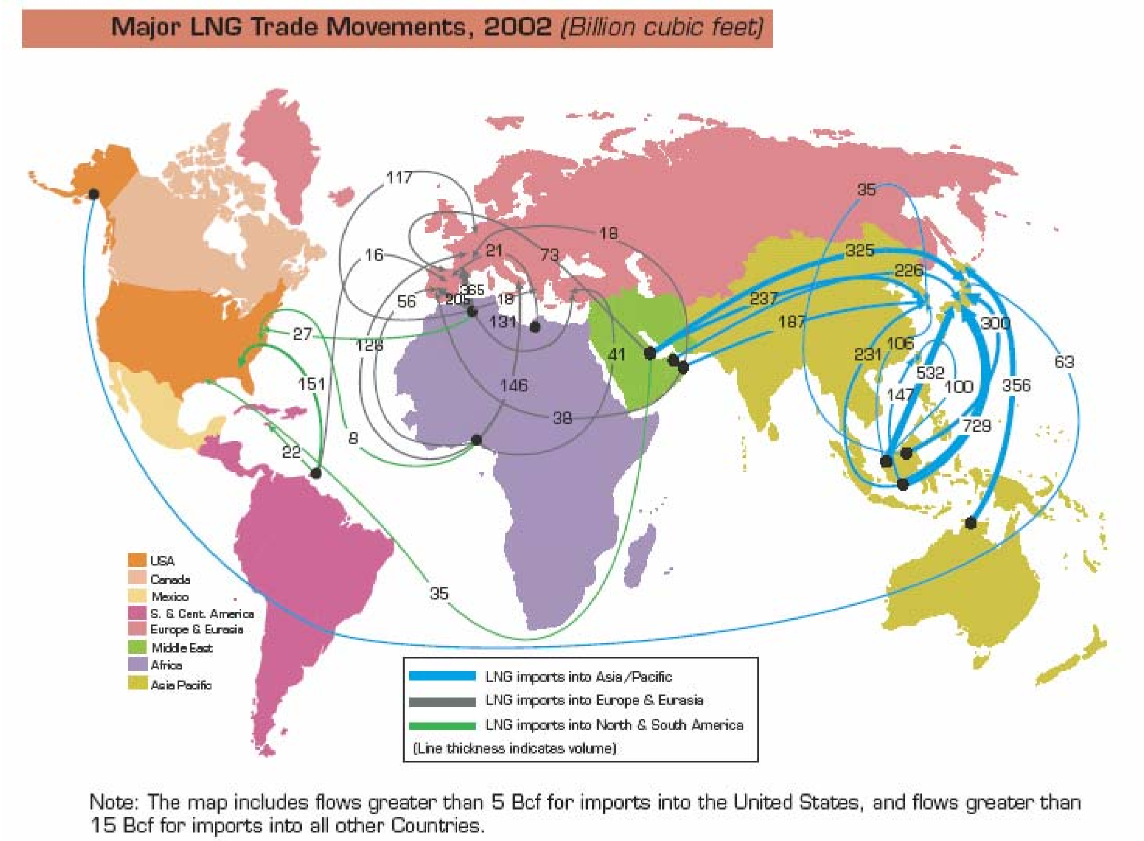

Figure 2.11 Major LNG Trade Movements (EIA[7], 2004)

The EIA estimates that the proposed Alaskan pipeline will be built in 2010[1]. This pipeline will bring natural gas from the Alaskan North Slope to Canada's MacKenzie River Valley, and tie into Canada's existing pipeline network. This pipeline is expected to increase imports of natural gas from Canada, till 2015[1]. After 2015, the increase demand and declining production will reduce the amount of natural gas imported from Canada[1].

The speed at which new LNG marine terminals will be built depends mainly on the regulation process. As mentioned before the FERC is the main regulatory agency for the onshore marine terminals. In 2002 the FERC took some steps to ease the regulatory process through the Hackberry Decision. This decision, in essence, made LNG terminals a supply point instead of a part of the transportation chain[2]. It was determined by FERC that since LNG was competitive with other sources of natural gas, it did not need further regulatory scrutiny. However, even though the FERC has eased regulations, there is the problem with overcoming the local and state regulation. Several LNG terminals have been proposed, but due to local opposition these proposals have been denied. One way companies are overcoming this obstacle is to build offshore LNG marine terminals.

Offshore marine terminals are regulated by the Coast Guard and MARAD. These terminals are less likely to be scrutinized by the local community because they are located several miles offshore. The Energy Bridge offshore marine terminal was the first new LNG terminal to be built in more than two decades. This project took a little over two years to permit. However, onshore marine terminals take longer to permit due to increased number of regulatory requirements compared with offshore terminals. Also, since LNG tankers are over 900ft long, it is easier for LNG ships to offload to terminals that are farther out in the ocean than this located on land. This is due to the fact that LNG ships have a large draft and need a wide area to maneuver. This requires onshore marine terminals to make sure that they are in deep enough water to accommodate the vessels. Also, LNG ships have to deal with local water traffic near the onshore marine terminal. In the case of the offshore marine terminal, the traffic near the terminal is very little which allows easy travel to and from the terminal.

Prices for LNG tankers have come down dramatically in the past 10 years. In 1995 a 138,000 cubic-meter capacity tanker cost $280 million. Today the same capacity tanker costs between $150 million to $160 million[7]. However with the recent increases in steel prices, costs of LNG tankers are likely to increase. This could lead to fewer LNG tankers being built as projects would become uneconomical due to the increase costs for LNG tankers.

CHAPTER III

NEURAL NETWORKS

3.1 Advantages/Disadvantages

Using neural networks to predict future behavior is a relatively new forecasting methodology. There are several advantages and disadvantages that neural networks offer in trying to forecast data.

The first advantage is that neural networks are very flexible. The flexibility of the neural network allows for modeling of linear and non-linear processes[10]. The neural network is therefore able to account for linear and non-linear parameters in the model. The second advantage of the neural network is the activation function, which provides robustness to the network[11]. Most neural networks use the sigmoid function, which reduces the effect of extremely large input values on the neural network. The third advantage of using neural networks is that neural networks "may work in situations where an explicit model-based approach fails."[11]. Linear or non-linear models by themselves can fail in forecasting, but neural networks have the ability combine these linearities and non-linearities in the same model. Finally, since neural networks are more flexible than other forecasting models, they can adapt easier to irregularities and unusual features in the data set.

There are a couple of inherent disadvantages when using neural networks. The first disadvantage is that neural networks need a large set of data points. Data points are used to train the data as well as test the data. If only a small amount of data points are available, then the neural network could "over fit" the data. In the event that the neural network is over fitted, any new data that is introduced could cause the neural network to become unstable. The other disadvantage of using neural networks is that the network does not provide an explicit model. Instead, neural networks is a black-box approach to forecasting, which does not provide much understanding of the data. However in explicit forecasting methods, the behavior of each input is known so the over-all behavior of the model is also known.

3.2 Current Uses

Neural networks have been used in several applications pertaining to the petroleum industry, and in other fields as well. The growing use of neural networks has established itself as an alternative to other forecasting techniques.

Neural networks have been used in the petroleum industry for a number of different applications. These applications can be split up into two main groups, one for predicting/characterizing properties, and the other for forecasting future supply/demand/prices. There have been several applications that fall under the predicting/characterization category and these include: well-testing[12], well logs[13], phase behavior[14], and naturally fractured reservoirs[15]. Under the second category, these applications include: natural gas supply[16], natural gas production[17], short-term natural gas prices[18], and oil and gas spot prices[19]. My thesis falls under the second category. Ultimately I will forecast future imports of natural gas into the U.S., specifically by way of LNG.

Neural networks have been used by other industries as well. In the financial sector, neural networks have been used in futures trading[20], and exchange rates[21]. In the environmental sector, neural networks have been used in ozone concentration and level[22], and in air quality[23]. From these examples we see that neural networks can be used in a number of different industries and areas.

3.3 Neural Network Study Model

Neural networks are modeled after the human brain. The human brain contains approximately one hundred billions neurons. These neurons are connected to as many as one thousand other neurons[24]. Each neuron receives many signals from other neurons, and when the combination of these signals exceed a certain threshold, or activation level, then the neuron fires sending signals to the other neurons. Neural networks try to mimic this structure and behavior after the human brain. Below is a simple flow diagram of what a neural network node "neuron" looks like.

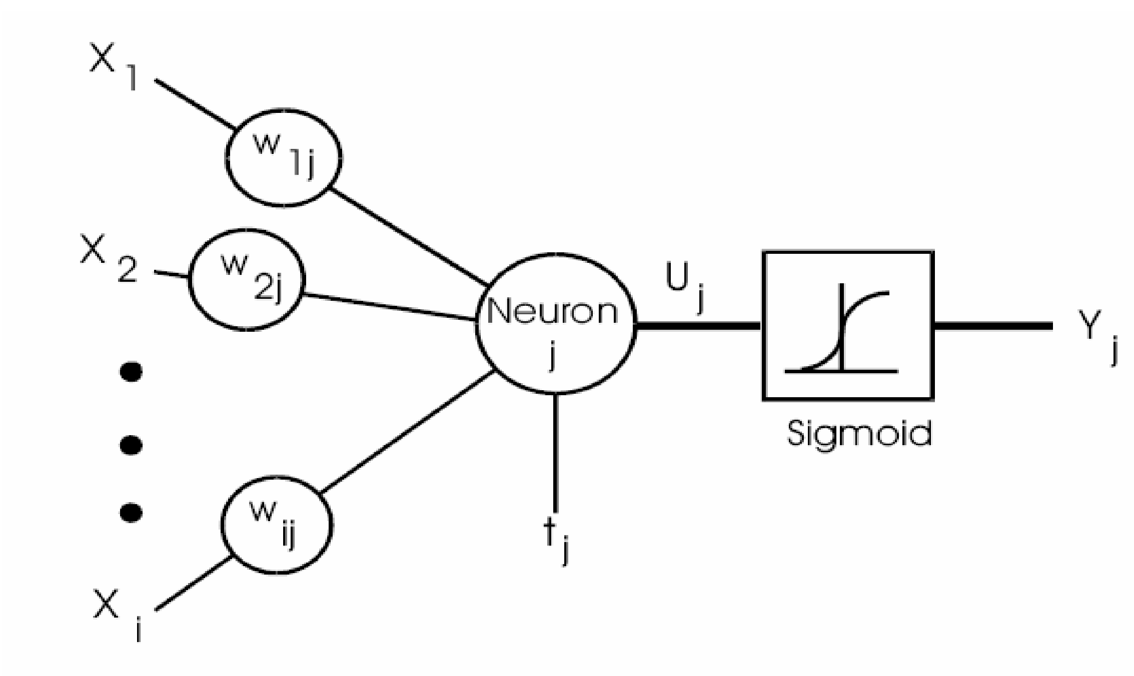

Figure 3.1 A Model Neuron (Shih[24], 1994)

In **Fig 3.1** we see that the neuron (node) is connected to inputs X_1, X_2, and X_i. The neuron receives these weighted inputs, which are a simply the inputs multiplied by the weights w_{1j}, w_{2j}, and w_{ij}. The neuron receives these weighted inputs and sums them up as defined by U_j. The sum of the weighted inputs is then modified by a previous threshold value, t_j, and then sent through the activation function. In this model the activation function is the sigmoid function. There are many activation functions to choose from such as Gaussian, hyperbolic, linear, and step. However the sigmoid activation function is the most common and will be used in this study. Once the modified sum of weighted inputs is sent through the activation function our result is Y_j.

In **Fig 3.2** we see a diagram of the sigmoid activation function. What the activation function does is take the output produced from the neuron, U_j+t_j, and gives it a value between zero and one. If the activation function receives a negative input then the resultant output will be inhibitory, and if the function receives a positive input then the output will be excitatory. This behavior is true for sigmoid, hyperbolic and linear activation functions.

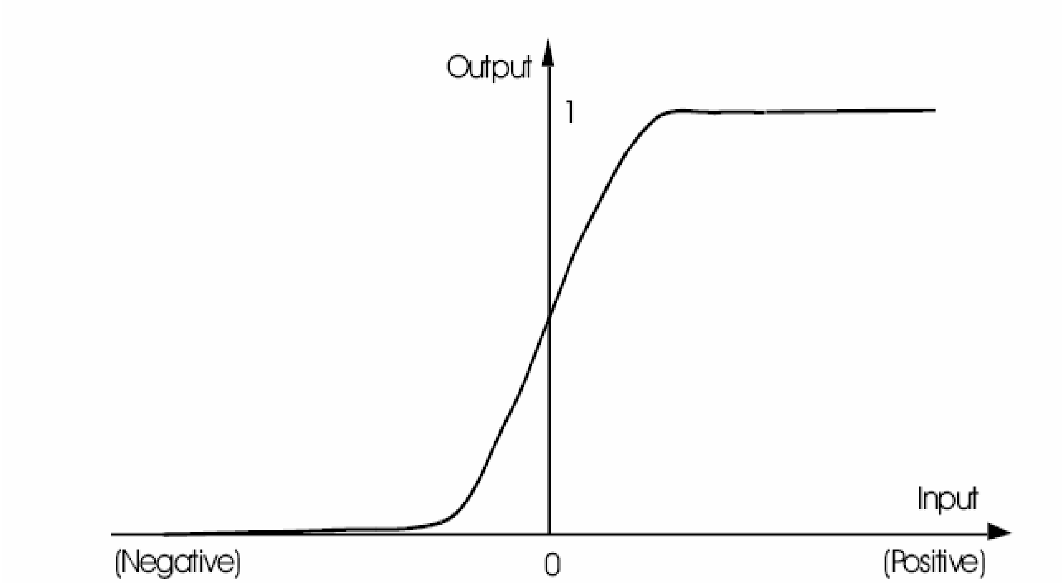

Figure 3.2 Sigmoid Activation Function (Shih[24], 1994)

When using neural networks to forecast data, the network first has to be trained. The way the network is trained is through a procedure called back propagation. **Fig. 3.3** shows the flow diagram of this procedure. As shown before, we see the neuron connected to the inputs and outputting a result by way of an activation function. When a neural network is training, the network takes this output and compares it to the actual value d_j. The difference between these two values is the error and is represented by e_j. This error is then back propagated through the network, specifically to the weights and the threshold value. Based on this error, the weights and the threshold value are adjusted accordingly.

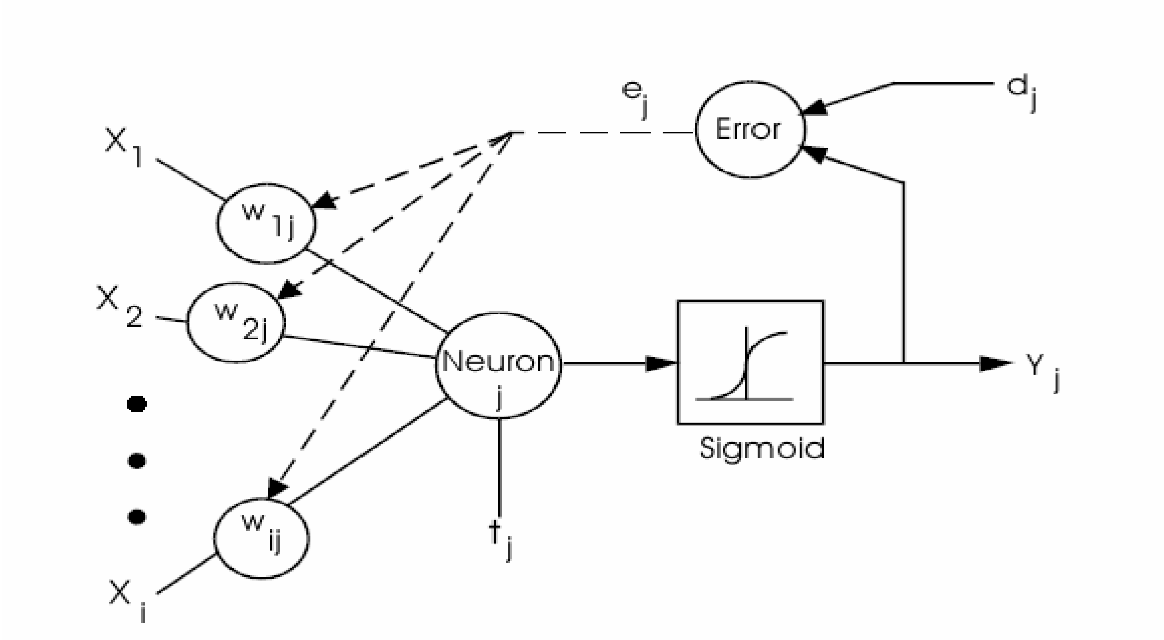

Figure 3.3 Diagram of a Back Propagation Network (Shih[24], 1994)

There are several equations that are used in a back propagation network. The first equation is used to adjust the weights of the output layer is as follows:

$$w_{ij} = w'_{ij} + LR * e_j * X_i$$

Where *i*, is the *i*th input and *j*, is the *j*th neuron. The new weight, w_{ij} is adjusted by the previous weight w'_{ij}, plus the product of the learning rate, LR, * error term, e_j, and the value of the input, X_i. The learning rate adjusts the magnitude of the correction term. A small value for LR will make learning slower, but more stable. A large value for LR will make learning faster, but more unstable. The error term, e_j, is computed by the following equation:

$$e_j = Y_j * (1-Y_j) * (d_j - Y_j)$$

Where Y_j is the output, $(1-Y_j)$ is Y_j compliment, and (d_j-Y_j) is the difference between the desired output, d_j, and the actual output. This error term equation is for the output layer.

If there are any hidden layers, the error term for these layers is defined by the following equation:

$$e_j = Y_j * (1-Y_j) * \sum(e_k * w'_{jk})$$

The terms e_j, Y_j, and $(1-Y_j)$ are the same as before, however the $(d_j - Y_j)$ term is replaced. The new term $\sum(e_k * w'_{jk})$ is the summation of next layer's error term times its weight. When the readjustment takes place, the output layer is computed first, and then each immediately proceeding layer is then computed, using the errors and weights from the succeeding layer.

Most neural networks consist of an input layer, one or more hidden layers, and an output layer. **Fig. 3.4** shows the overall diagram of a common back propagation network. In a back propagation network, the input layer takes in the inputs to be used for the neural network. These input layer nodes then distribute its signals to the hidden layer. The hidden layer nodes take these signals and attempt to categorize or detect the features of these signals[24]. The output layer then collects these features and produces an output.

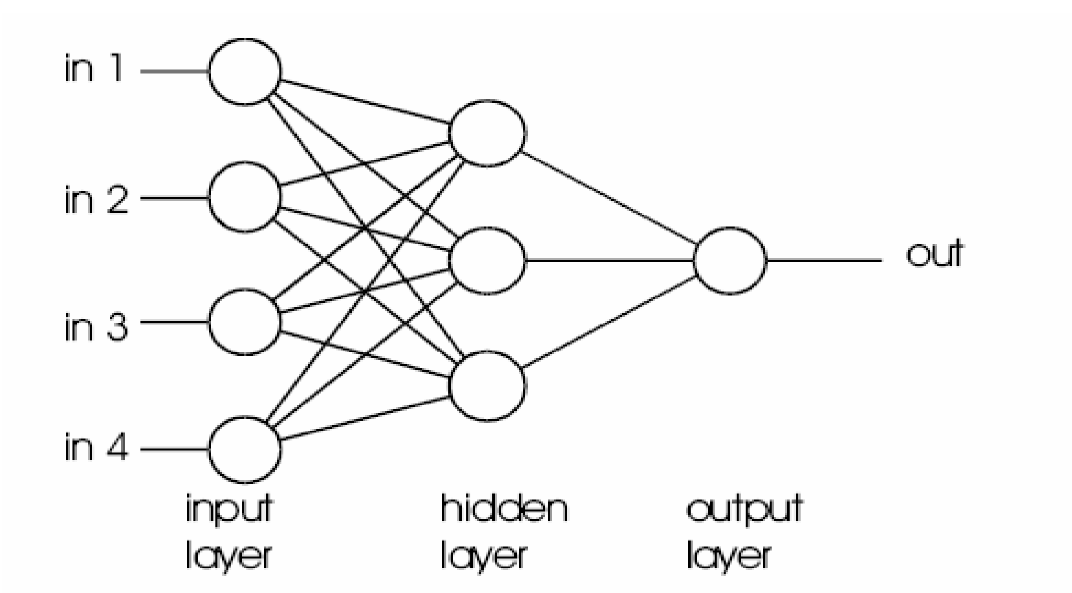

Figure 3.4 Back Propagation Network with One Hidden Layer (Shih[24], 1994)

Neural networks can vary greatly in size and structure. Most common neural networks have only one hidden layer. However, a neural network can have many hidden layers, and different number of hidden nodes per layer. There are many neural network software available to construct these networks. However, the art is in choosing the right inputs and structure to accurately model the behavior of past data and to predict future behavior.

CHAPTER IV

ORGANIZATIONAL METHODOLOGY

4.1 Data Gathering

A number of steps were taken to preprocess the data to run the data through the neural network simulation. First, I gathered data on various inputs I thought would be important in factors in forecasting LNG imports. Next, I normalized the data all the data points. Thirdly, I organized them into different groups based on their similarity. Finally, I used a correlation matrix to make sure that there wasn't a redundant factor in any of the groups.

Since LNG imports have been relatively flat for 30 years, I chose a wide variety of factors that I believed might have some influence on LNG imports. In choosing these factors, I wanted to be sure that the data collected on these factors were consistent. And since I was going to ultimately compare my forecast with the EIA forecast, I decided to choose factors for which I could obtain data from the EIA website. After examining and studying the information available on the website I chose the following factors:

- Natural Gas Consumption (TCF)
- Natural Gas Production (TCF)
- Natural Gas Wellhead Price
- Crude Oil Wellhead Price
- Coal Electric Generation (BKWH)
- Petroleum Electric Generation (BKWH)
- Natural Gas Electric Generation (BKWH)
- Nuclear Electric Generation (BKWH)

- Renewable Electric Generation (BKWH)
- Crude Oil $/MMBTU
- Natural Gas $/MMBTU
- Coal $/MMBTU
- Electricity $/MMBTU
- Population
- GDP (Year 2000 dollars)
- Residential Consumption of Natural Gas (TCF)
- Commercial Consumption of Natural Gas (TCF)
- Industrial Consumption of Natural Gas (TCF)
- Electrical Generation Consumption of Natural Gas (TCF)
- Residential Energy Consumption of Natural Gas (Quad BTU)
- Commercial Energy Consumption of Natural Gas (Quad BTU)
- Industrial Energy Consumption of Natural Gas (Quad BTU)
- Electrical Generation Energy Consumption of Natural Gas (Quad BTU)

For historical data I used the EIA Annual Energy Review 2003 for data from 1970-2003. The reason why 1970 is the starting point for the data is because the U.S. received its first LNG shipment from Algeria in that year. For the forecasted input data I used EIA Annual Energy Outlook 2005 for data from 2004 through 2025.

4.2 Normalization

Once all the data was collected I normalized all the data using the mean standard deviation method. This is the most common way from normalizing data to be used in neural networks. The formula for this equation is as follows:

$$X'_i = \frac{(X_i - \mu_i)}{\sigma_i}$$

There is some debate in whether normalization is necessary. For example, some researchers (Gorr[25], 1994) believe that preprocessing data has no affect on the neural network, because it is able to capture all of the underlying patterns[10]. However, empirical studies (Nelson et al.[26], 1999) find that pre-processing data is critical for improving forecasting performance[10]. Also, Zhang and Qi (2002)[27] find that for time series forecasting containing trend and seasonal variations, data preprocessing "should be the most appropriate way to build neural networks for best forecasting performance"[10].

4.3 Organization of Factors

Looking at the input factors I noticed that they could be broken into groups. The number of inputs in each group was dependent on two criteria. This first criterion was that the group had to have at least three inputs. Using only two inputs would increase the neural networks dependence on a single factor. The second criterion that I used was that the neural network could have no more than five inputs. The reason why I limited the number of inputs to five is that I ran the risk of over fitting the neural network. An over fitted neural network has close to as many weights as it does training data points. A characteristic of an over fitted network is that when new data is introduced the new output can be highly scattered.

Taking these two criteria I then grouped the inputs by similarity. For example one group that I used was the electricity generation by fuel. In this group we have the following inputs: coal, petroleum, natural gas, nuclear, and renewable electric generation. The other groups that I used are as follows: residential, commercial, industrial, and electrical consumption of natural gas on a TCF basis; residential, commercial, industrial, and electrical consumption of natural gas on a Quad BTU basis; and $/MMBTU of crude oil, natural gas, coal, and electricity.

4.4 Correlation Matrix

Since all the groups have been organized, a correlation matrix was run on each group to determine if there were any redundant factors. I made a rule that if the correlation coefficient for the factor was greater than 0.95, then I would eliminate it from the model. For the $/MMBTU group the correlation matrix is displayed in **Table 4.1**.

Table 4.1 Correlation Matrix for $/MMBTU Model

Factor ($/MMBTU)	Crude Oil	Natural Gas	Coal	Electricity
Crude Oil	1			
Natural Gas	0.7998	1		
Coal	0.7328	0.6826	1	
Electricity	0.6751	0.9279	0.7964	1

The highest correlation coefficient is between natural gas and electricity with a correlation coefficient of 0.9279. Since the correlation coefficient is below 0.95, both factors will remain in the $/MMBTU model. The next group that was looked at was the Natural Gas Energy Consumption. **Table 4.2** displays the results from the correlation matrix.

Table 4.2 Correlation Matrix for Natural Gas Energy Consumption Model

Factor (Energy Consumption)	Residential	Commercial	Industrial	Electrical Generation
Residential	1			
Commercial	0.5382	1		
Industrial	0.6157	0.4046	1	
Electrical Generation	0.4787	0.7643	0.4095	1

None of the factors for this model are highly correlated, thus all the factors are included in the model. The next group is the Natural Gas Consumption model. The results of the correlation matrix are displayed in **Table 4.3**.

Table 4.3 Matrix for Natural Gas Consumption Model

Factor (Consumption)	Residential	Commercial	Industrial	Electrical Generation
Residential	1			
Commercial	0.5263	1		
Industrial	0.6095	0.5426	1	
Electrical Generation	0.4486	0.7675	0.5503	1

In this correlation matrix, correlations ranged between 0.4486 and 0.7675. Overall, none of the factors were highly correlated with the others, so all factors were used in this model. The Energy Stack model's correlation matrix is displayed in **Table 4.4**.

Table 4.4 Correlation Matrix for Energy Stack Model

Factor (Energy Stack)	Coal	Petroleum	Natural Gas	Nuclear	Renewable
Coal	1				
Petroleum	-0.8055	1			
Natural Gas	0.7152	-0.4775	1		
Nuclear	0.9863	-0.7571	0.7317	1	
Renewable	0.6933	-0.6218	0.5449	0.6681	1

This correlation shows that some factors are positively correlated and negatively correlated to other factors. The highest correlation was between nuclear and coal with a value of 0.9863. Since this value is over 0.95, I had to choose which factor to eliminate. Since coal electric generation provides more electricity in the U.S. than nuclear electric generation, I chose to eliminate nuclear electric generation from the energy stack model.

The overall purpose of creating these correlation matrixes was to determine if there were any redundant factors. It was found in the energy stack model that nuclear electric generation and coal electric generation were highly correlated, and thus were redundant factors. In all the other models, it was determined that none of the factors were redundant.

CHAPTER V

STANDARD APPROACH

5.1 Introduction

The standard approach of reducing the number of input variables is somewhat similar to grouping approach. The data collection steps and normalization of the factors are the same, but determining which of these inputs to use is different.

5.2 Screening Variables

With the standard approach, the first step in screening the input variables is to create a correlation matrix to determine which factors have the highest linear correlation to the output. The output in this case is U.S. LNG imports. The correlation coefficients with respect to the output are shown in **Table 5.1**.

Table 5.1 Correlation Coefficients for All Inputs Compared to the Output, U.S. LNG Imports

Input Factor	Correlation Coefficient
Consumption (TCF)	0.2567
Production (TCF)	-0.0047
Natural Gas Wellhead Price	0.8043
Crude Oil Wellhead Price	0.2405
Coal Electric Generation (BKWH)	0.5362
Petroleum Electric Generation (BKWH)	-0.2509
Natural Gas Electric Generation (BKWH)	0.6850
Nuclear Electric Generation (BKWH)	0.5414
Renewable Electric Generation (BKWH)	0.2528
Crude Oil $/MMBTU	0.5598
Natural Gas $/MMBTU	0.6681
Coal $/MMBTU	0.1404
Electricity $/MMBTU	0.4130
Population	0.6386
GDP	0.6705
Residential Consumption of Natural Gas (TCF)	0.2723
Commercial Consumption of Natural Gas (TCF)	0.5793
Industrial Consumption of Natural Gas (TCF)	-0.0193
Electrical Generation Consumption of Natural Gas (TCF)	0.6514
Residential Energy Consumption of Natural Gas (QUAD BTU)	0.2905
Commercial Energy Consumption of Natural Gas (QUAD BTU)	0.5796
Industrial Energy Consumption of Natural Gas (QUAD BTU)	-0.1822
Electrical Generation Energy Consumption of Natural Gas (QUAD BTU)	0.6524

The ten most correlated inputs are shown below in **Table 5.2**, in descending order. Natural gas wellhead price has the highest correlation factor with a correlation coefficient of 0.8043.

Table 5.2 Ten Most Correlated Factors Arranged in Descending Order

Input Factor	Correlation Coefficient
Natural Gas Wellhead Price	0.8043
Natural Gas Electric Generation (BKWH)	0.6850
GDP	0.6705
Natural Gas $/MMBTU	0.6681
Electrical Generation Energy Consumption of Natural Gas (QUAD BTU)	0.6524
Electrical Generation Consumption of Natural Gas (TCF)	0.6514
Population	0.6386
Commercial Energy Consumption of Natural Gas (QUAD BTU)	0.5796
Commercial Consumption of Natural Gas (TCF)	0.5793
Crude Oil $/MMBTU	0.5598

These inputs were then screened again by using a correlation matrix to determine if there were any redundant factors. From this correlation matrix I found that several factors were redundant. **Table 5.3** lists the redundant factors and their correlation coefficients.

Table 5.3 Redundant Factors and Their Correlation Coefficients

Redundant Factors	Correlation Coefficient
Natural Gas Wellhead Price/ Natural Gas $/MMBTU	0.9547
Natural Gas Electrical Generation/ Electrical Generation Energy Consumption of Natural Gas	0.9513
Natural Gas Electrical Generation/ Electrical Generation Consumption of Natural Gas	0.9550
GDP/ Population	0.9947
Commercial Energy Consumption of Natural Gas/ Commercial Consumption of Natural Gas	0.9989
Electrical Energy Consumption of Natural Gas/ Electrical Consumption of Natural Gas	0.9997

Since I had to eliminate one of the two redundant factors in each group, I decided to eliminate the factor that had the lower correlation coefficient compared to LNG imports. The following factors were eliminated: Natural Gas $/MMBTU; Electrical Generation Energy Consumption of Natural Gas; Electrical Generation Consumption of Natural Gas; Population; and Commercial Consumption of Natural Gas. That leaves the following factors to be used in the neural network simulation: Natural Gas Wellhead Price; Natural Gas Electric Generation; GDP; Commercial Energy Consumption of Natural Gas; and Crude Oil $/MMBTU.

5.3 Neural Network Setup

Before running the neural network several steps have to be taken to setup the neural network. The first step is to specify how many hidden layers and how many nodes per hidden layer in the network. Since there are a limited amount of data points, I chose to have only one hidden layer. Having hidden layers allows for more flexible and complicated models to be fitted[11]. I originally planned to use three hidden nodes in the hidden layer. However the number of weights used in this model is 24, which is close to the 33 data points that I had to test with. I then decided to use two hidden nodes which resulted in 16 weights for all the models. The second step is training the neural network. The data points from 1970-2003 are the data points that I trained with. To train the network some of the data points are designated TRAIN and the other data points are designated as TEST. There are several splitting percentages that are suggested in literature. The most common splitting percentages are 70%/30%, 80%/20%, and 90%/10%. Which splitting percentages that are used is not as important as making sure that enough data points are available for learning, validation and testing. Granger

(1993)[28] suggests that for non-linear modeling, 20% of the data points should be used for testing[10]. Hoptroff (1993)[29] recommends at least 10 data be used, and Ashley (2003)[30] suggests using a much larger number of data points for testing. In my simulation I used 80% of the data for training and 20% of the data for testing.

CHAPTER VI

RESULTS

6.1 $/MMBTU Model

In this model I used the $/MMBTU of four different commodities: coal, natural gas, crude oil, and electricity. This is the first model tested in the neural network. During the training phase I achieved a RMS error of 0.08039 TCF^2 after 49,924 training epochs. A training epoch is complete when the all the defined training cases have been processed. The reason why I ran a large number of training epochs is to make sure that the neural network model had stabilized during the training. **Fig. 6.1** shows the learning rate for the neural network. From this figure we can see that most of the learning occurred early on and then flattened indicating that little or no learning was taking place. **Table 6.1** shows the result from training/testing phase for the first model.

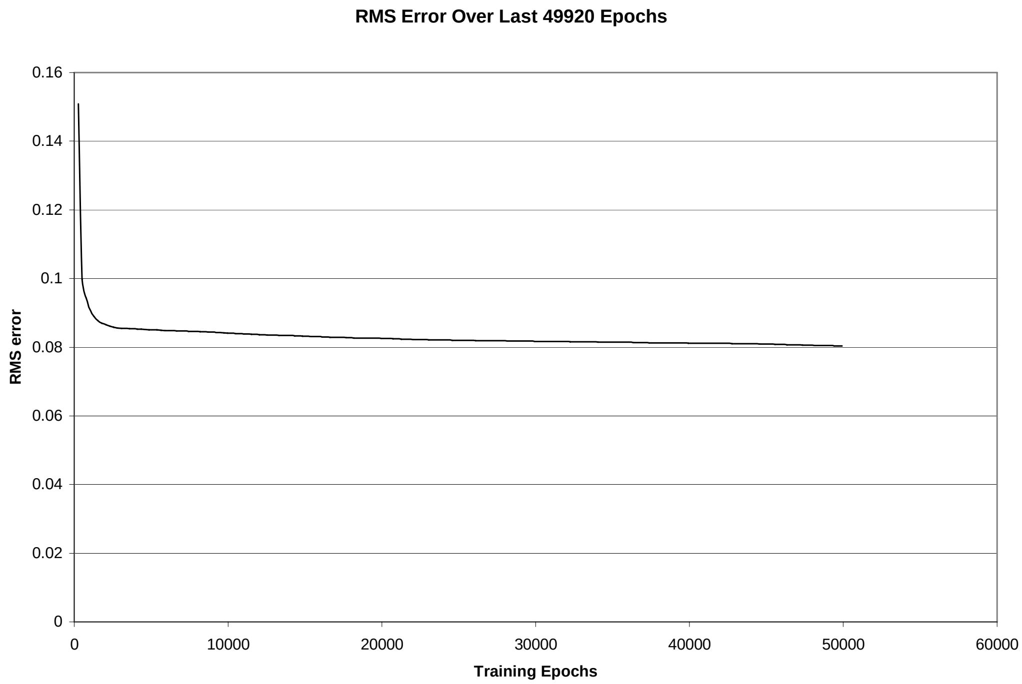

Figure 6.1 RMS Training Error Over the Last 49920 Training Epochs

Table 6.1 Training and Testing Error for the $/MMBTU Model

Training		Testing	
RMS (TCF^2)	Correct %	RMS (TCF^2)	Correct %
0.08039	93	0.1720	100

After I trained the network I then used the network parameters and forecasted future LNG Imports based on the forecasted input data. **Fig 6.2** shows the neural network prediction compared to the EIA forecast.

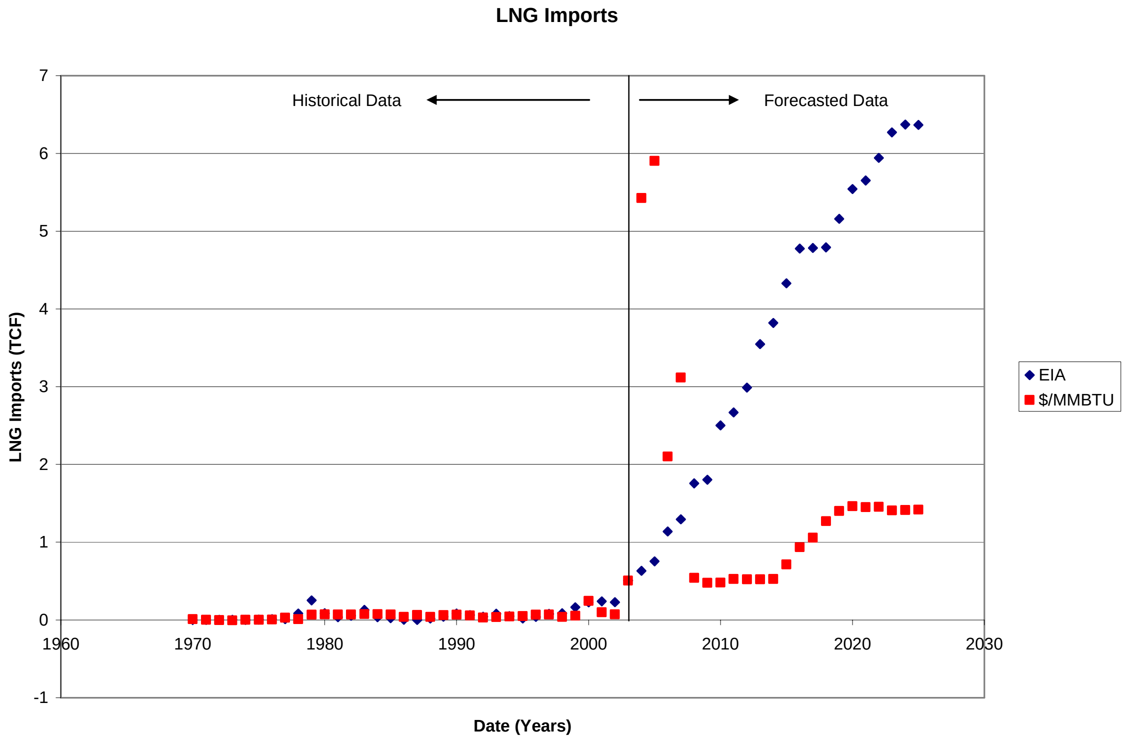

Figure 6.2 Results for the $/MMBTU Model Compared to the EIA Forecast

From this figure we can see a couple of interesting features. First, the first four data points for the predicted data are highly scattered. One possible explanation is that the neural network is over fitted. However I took steps to resolve this issue and do not believe this is the case. The second feature is that maximum expected amount of LNG imports is 1.5 TCF which is considerably lower than the EIA forecast. Due to the initial forecasted data points being unstable, this scenario is not very useful.

6.2 Natural Gas Energy Consumption Model

In this model I used four factors: residential, commercial, industrial, and electrical energy consumption of natural gas on a Quad BTU basis. The results for the training/testing part of the simulation are listed below in **Table 6.2**.

Table 6.2 Training and Testing Error for the Natural Gas Energy Consumption Model

Training		Testing	
RMS (TCF^2)	Correct %	RMS (TCF^2)	Correct %
0.04304	93	0.2104	86

Again I ran a large number of training epochs to makes sure the neural network model was stabilized. In this scenario the forecasted data points turned out to be smoother than the last scenario. The results were plotted against the EIA forecast and the results are shown below in **Fig. 6.3**.

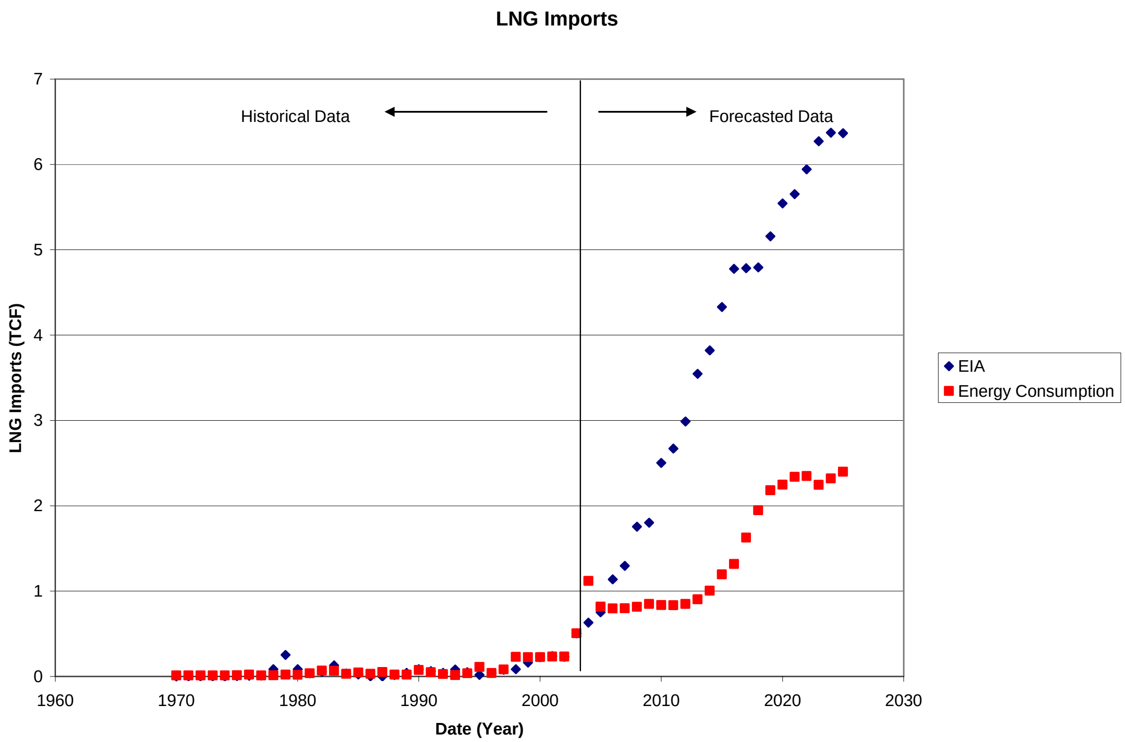

Figure 6.3 Results for the Natural Gas Energy Consumption Model Compared to the EIA Forecast

As shown, the forecast trend of the energy consumption model is a lot smoother than that of the previous model. Again there is a relatively low maximum amount of

LNG imports compared with the EIA forecast. This model predicts that LNG imports will max out at about 2.5 TCF.

6.3 Natural Gas Consumption Model

In this model I used for factors: residential, commercial, industrial, and electric consumption of natural gas on a TCF basis. The results from the training/testing portion of the simulation are listed in **Table 6.3**.

Table 6.3 Training and Testing Error for the Natural Gas Consumption Model

Training		Testing	
RMS (TCF^2)	Correct %	RMS (TCF^2)	Correct %
0.04491	96	0.1102	100

The results from the simulation are shown below in **Fig. 6.4**. From this graph we see some interesting results. In this model there is a smooth trend compared to that of the first model. Secondly we see that the amount of LNG imports peaks out at a value of about 6.6 TCF. Surprisingly, this value is close to the maximum amount of LNG imports that the EIA forecasts. However, I am skeptical of the rapid rise in LNG imports from 2008 till 2012. I believe that that rapid of a rise is possible, but not at such an early date, because most terminals will likely be built during that time and come online in 2012. Perhaps, the forecast would be more realistic if the trend was shifted 3-4 years to the right.

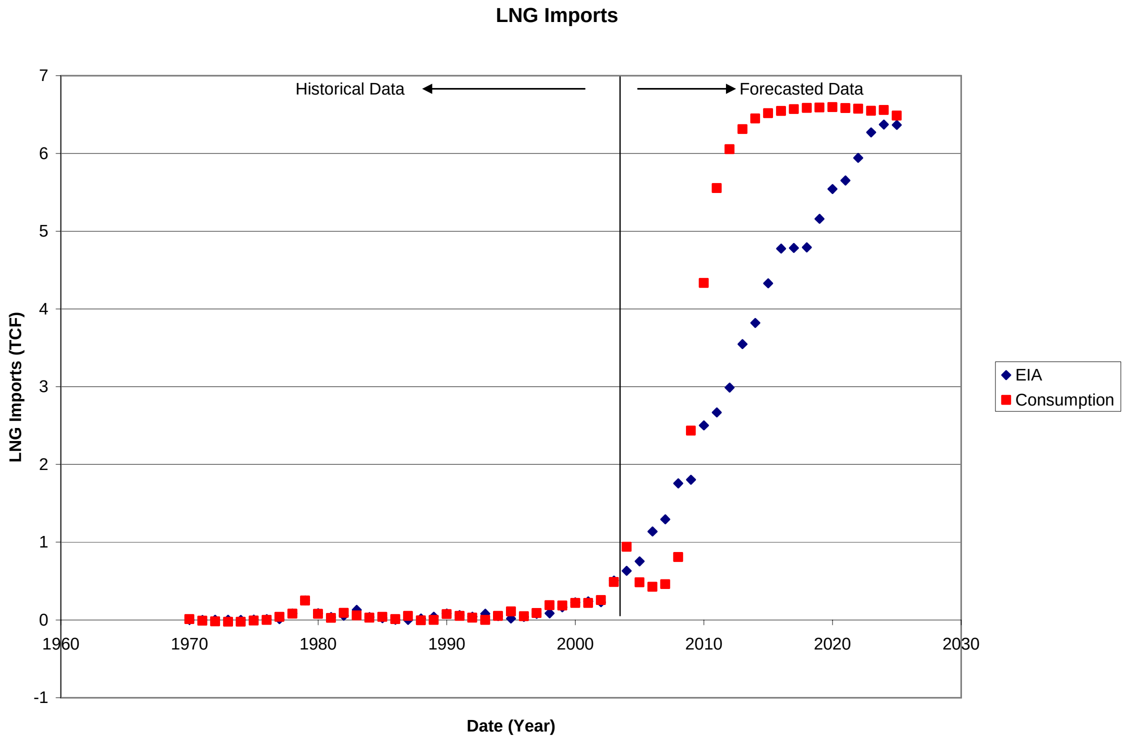

Figure 6.4 Results for the Natural Gas Consumption Model Compared to the EIA Forecast

6.4 Energy Stack Model

This model is based off of how utility companies use different sources of power to provide electricity to its consumers. Basically, utilities use power sources such as hydroelectric, nuclear, and coal as their base supply of power. Utilities use electricity generated from natural gas power plants as a peak-shaving plant. A peak-shaving plant is used when there is a heavy demand for electricity, such as on a hot summer day. The factors I used in this model are as follows: coal, petroleum, natural gas, and renewable electric generation on a BKWH basis. The results from the training/testing portion of the simulation are listed in **Table 6.4**.

Table 6.4 Training and Testing Error for the Energy Stack Model

Training		Testing	
RMS (TCF^2)	Correct %	RMS (TCF^2)	Correct %
0.0889	93	0.1859	86

Fig. 6.5 shows the results from the energy stack model. This model represented close to what the U.S. might see as far as LNG imports. First, the trend of the forecasted data is the smoothest of all models studied. Secondly, the maximum amount of LNG imports is about 6.6 TCF which is close to the EIA forecast. Also, we see a big rise in LNG imports starting in 2013, which I think is most likely time frame given that the permitting process takes at least two years, and it takes another few years until the first shipment of LNG arrives for a given project. In **Fig. 6.4** we began to see the rapid rise in LNG imports beginning in 2008, which I believe is too early. Another feature of this trend is that there is an exponential rise in LNG imports where in the EIA forecast, there seems to be almost a linear trend in the amount of LNG imports. I think an exponential trend is more likely because there has been a very large increase in the amount of permit applications filed. Even if a fraction of these proposed projects come to fruition, I think that these new terminals are going to add a significant amount of LNG terminal capacity in a short amount of time.

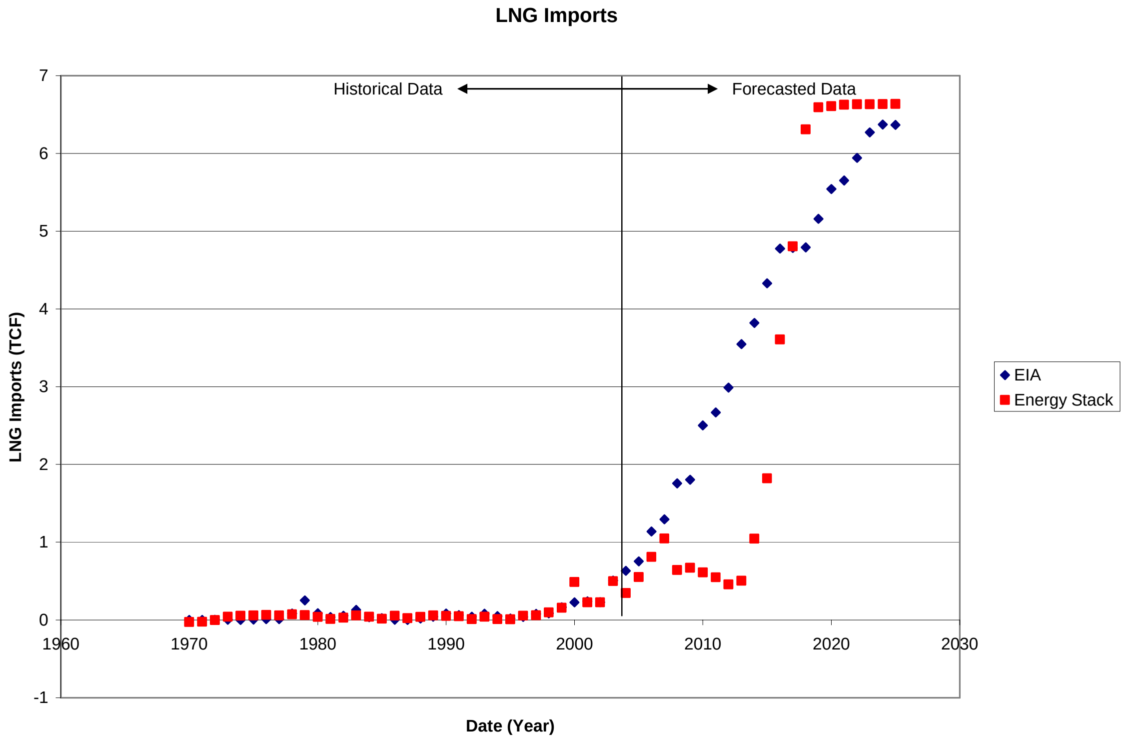

Figure 6.5 Results for the Energy Stack Model Compared to the EIA Forecast

6.5 LNG Terminal Capacity Model

This model was created to compare the neural network forecasts to a forecast based on the 2004 terminal capacity of the four LNG terminals in the U.S. For this model I took the amount of LNG imported in 2004 and divided it by four to get an average amount of LNG imported per terminal. From there I made two forecasts based on different assumptions. The first forecast was based on adding one equivalent terminal per year, and the second forecast was based on adding two equivalent terminals per year. I assumed the maximum amount of equivalent terminals was 18 based on the current number of applications filed. The results are shown in **Fig. 6.6**.

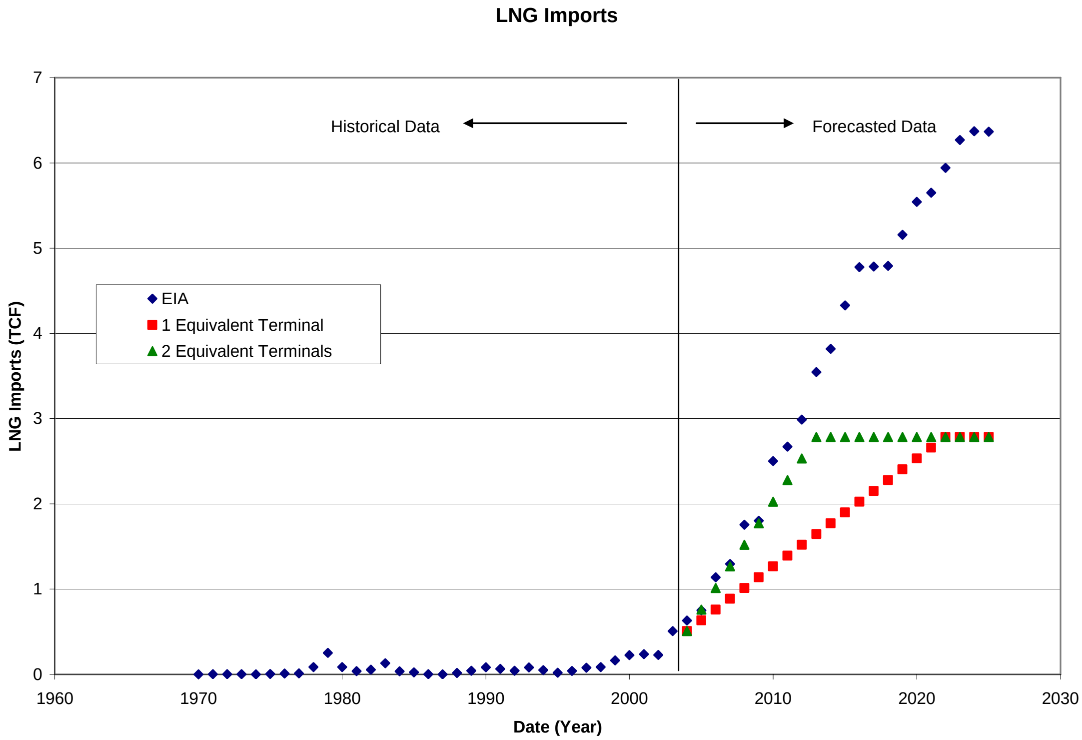

Figure 6.6 Results for the LNG Terminal Model Compared to the EIA Forecast

Several interesting observations can be drawn from this figure. First, the two terminals per year equivalent seems to follow the trend of the EIA forecast. If the number of maximum equivalent terminals were increased, this would result in the two terminals per year model paralleling the EIA forecast. However, this model is flawed, because the existing four terminals have filed permits to increase their capacity. Also most of the proposed terminals have capacity equal to or greater than the capacity of the four existing LNG terminals.

6.6 Standard Methodology Results

This neural network simulation used the following factors: Natural Gas Wellhead Price; Natural Gas Electric Generation; GDP; Commercial Energy Consumption of

Natural Gas; and Crude Oil $/MMBTU. The results from the training/testing phase of the simulation are presented in **Table 6.5**.

Table 6.5 Training and Testing Error for The Standard Model

Training		Testing	
RMS (TCF^2)	Correct %	RMS (TCF^2)	Correct %
0.0760	93	0.1487	100

Fig. 6.7 shows the results from using the standard approach. There are a few interesting features that we can see from this figure. The first few forecasted data points don't seem to follow the overall trend. This feature is also present in **Figures 6.2**, **6.4**, and **6.5**. This could be due to the fact that seems to be some disconnect between the 2003 historical data and the 2004 forecasted data. Also from this figure we see that the maximum amount of U.S. LNG imports predicted is 2 TCF. This figure is about 3 times less than what the EIA has forecasted. Given the number of permits filed and the planned expansion for the four current onshore marine terminals, I think this model is very pessimistic.

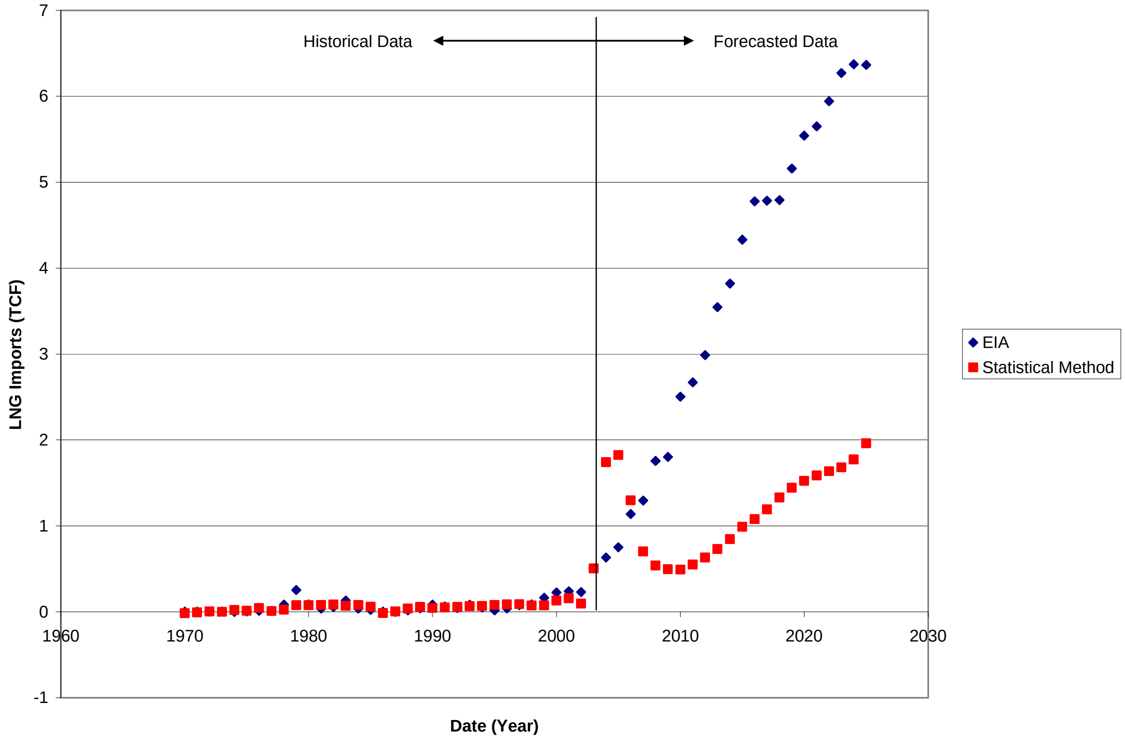

Figure 6.7 Results for the Standard Model Compared to the EIA Forecast

6.7 Forecast Comparisons

Below are **Table 6.6** and **Fig. 6.8**, which summarizes the results from all five scenarios. From this table we see that the $/MMBTU model, Energy Stack model and the Standard model all had RMS errors about twice as high as that of the Energy Consumption and Consumption model. This error could have been due to the limitation on the number of weights in the neural network. For the remaining two scenarios the RMS error for the training and testing portion were generally close to one another.

Table 6.6 Training and Testing Error for All Models

Scenario	Training		Testing	
	RMS (TCF^2)	Correct %	RMS (TCF^2)	Correct %
$/MMBTU	0.08039	93	0.1720	100
Energy Consumption	0.04304	93	0.2104	86
Consumption	0.04491	96	0.1102	100
Energy Stack	0.08898	93	0.1859	86
Standard	0.07603	93	0.1487	100

In **Fig. 6.8** shows the combined results from all the scenarios. The energy stack and consumption model predict a rapid increase in LNG imports. However, as pointed out before, the energy stack model is more realistic in that its rapid increase takes place 3-4 years later than the consumption model. Again, **Fig. 6.8** shows the instability of the $/MMBTU model in the first four data points for the forecasted data. The last observation is that the two equivalent terminal/year models follow the EIA forecast's linear trend.

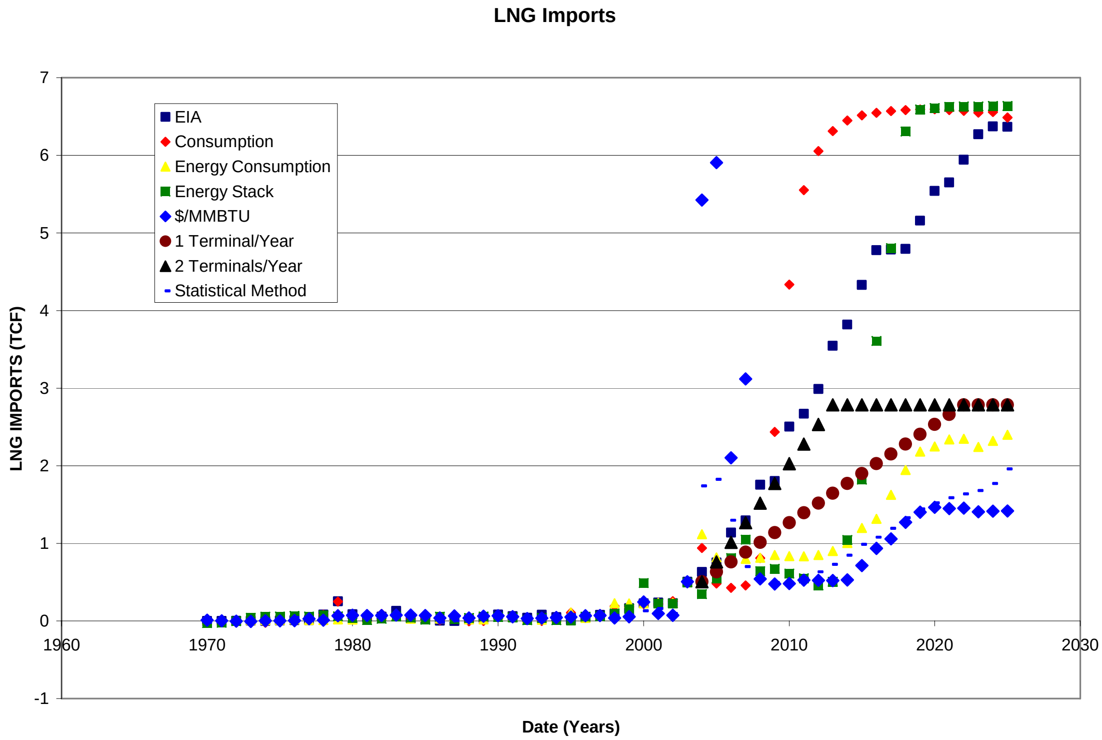

Figure 6.8 Results from All Models Compared to the EIA Forecast

CHAPTER VII

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

7.1 Summary

Even though the U.S. has been importing LNG for some time now, it has been only recently that there has been a renewed interest. The U.S. is on the brink of importing several TCFs of LNG, to make up the supply/demand gap. This study attempted to provide a brief overview of the LNG market, not only in the U.S. but abroad as well. Also the study provided some insight into the natural gas market and some of the economical and political factors affecting current and future LNG projects.

This study developed five different neural network models to predict LNG imports into the U.S. Two out of the five models proved useful in identifying general trends that the U.S. might expect to see in the coming years. This study provided an alternate way to forecast LNG imports.

7.2 Conclusions

1. Developed a unique methodology by organizing factors into groups based on similarity.
2. The energy stack model and the consumption of natural gas model forecasted a non-linear trend in U.S. LNG imports, compared to the linear trend forecasted by the EIA.
3. The energy stack model and consumption of natural gas model both forecasted U.S. LNG imports in 2025 to be about 6.5 TCF.

4. The standard model, $/MMBTU, and energy consumption of natural gas model all gave considerably lower forecasts when compared to the EIA forecast.
5. The energy stack model was the most realistic model, due to the non-linear trend in the forecast, when the rapid increase of LNG imports occurred, and the amount of U.S. LNG imports predicted in 2025.

7.3 Recommendations

The only recommendation that I have for this study is the need for more data. If more data was available then different neural network models could be run with more inputs. Due to the lack of data I was limited to a neural network model with a maximum of five inputs, and two hidden nodes in the hidden layer. If I used a greater number of inputs or hidden nodes, I ran the risk of over fitting the neural network.

REFERENCES

1. "Annual Energy Outlook 2005," DOE/EIA-0383, Office of Integrated Analysis and Forecasting, U.S. Dept. of Energy, EIA, Washington DC (2005).

2. Energy Information Agency, Internet Homepage: http://www.eia.doe.gov/ [Accessed April 27, 2005].

3. "Natural Gas Transmission and Distribution Model of the National Energy Modeling System," DOE/EIA-M062, Office of Integrated Analysis and Forecasting, U.S. Dept. of Energy, EIA, Washington DC (2004).

4. Tobin, J.: "Natural Gas Transportation – Infrastructure Issues and Operational Trends," Natural Gas Division, U.S. Dept. of Energy, EIA, Washington DC (2001), Internet Home Page: http://tonto.eia.doe.gov/dnav/ng/ng_sum_analysis.asp.

5. Steffy, L.: "Problems at Pemex Flow Deeper Than Management," *Houston Chronicle (November 2004).*

6. Greenwald, G.B.: *Liquefied Natural Gas: Developing and Financing International Energy Projects*, Kluwer Law International., Cambridge, MA (1998)

7. "The Global Liquefied Natural Gas Market: Status & Outlook," DOE/EIA-0637, Energy Information Agency, U.S. Dept. of Energy, Washington DC (2003).

8. "U.S. LNG Markets and Uses: June 2004 Update," Energy Information Agency, U.S. Dept. of Energy, Washington DC (2004)

9. Eisbrenner, K. (2004) *Excelerate Energy*. Society of Petroleum Engineers. Houston Petroleum Club. Presented information on Excelerate's Energy Bridge Vessel. Houston, TX, 9 September 2004.

10. Zhang, G.P.: *Neural Networks in Business Forecasting*, IRM Press, Hershey, PA (2004).

11. Hyndman, R.J., Makridakis, S., and Wheelwright, S.C.: *Forecasting Methods and Applications*, John Wiley & Sons, Hoboken, NJ (1998).

12. Al-Kaabi, A.U. and Lee, W.J.: "Using Artificial Neural Nets to Identify the Well-Test Interpretation Model," *SPEFE*, (September 1993), 233.

13. Mohaghegh, S.: "Virtual-Intelligence Applications in Petroleum Engineering: Part I-Artificial Neural Networks," *JPT*, (September 2000), 64.

14. Habiballah, W.A., Startzman, R.A., and Barrufet, M.A.: "Use of Neural Networks for Prediction of Vapor/Liquid Equilibrium KValues for Light-Hydrocarbon Mixtures," *SPERE*, (May 1996), 121.

15. El Ouahed, A.K., Jokhio, S.A., Mazouzi, A., and Tiab, D.: "Application of Artificial Intelligence to Characterize Naturally Fractured Reservoirs," paper SPE 84870 presented at the 2003 SPE International Improved Oil Recovery Conference. Kuala Lumpur, Malaysia, 20-21 October.

16. Al-Fattah, S.M., and Startzman, R.A.: "Forecasting World Natural Gas Supply," *Journal of Petroleum Technology*, (2000) **52**, 62-72.

17. Al-Fattah, S.M., and Startzman, R.A.: "Neural Network Approach Predicts U.S. Natural Gas Production," *SPE Production and Facilities*, (2003) **18**, 84-91.

18. Reiter, D.F., and Economides, M.J.: "Prediction of Short-term Natural Gas Prices Using Econometric and Neural Network Models," paper SPE 52960 presented at the 1999 SPE Hydrocarbon Economics and Evaluation Symposium. Dallas, Texas, 21-23 March.

19. Agbon, I.S., and Araque, J.C.: "Predicting Oil and Gas Spot Prices Using Chaos Time Series Analysis and Fuzzy Neural Network Model," paper SPE 82014 presented at the SPE Hydrocarbon Economics and Evaluation Symposium. Dallas, Texas, 5-8 April.

20. Kaastra, I., and Boyd, M.S.: "Forecasting Futures Trading Volume Using Neural Networks," *The Journal of Futures Markets*, 1995 **15(8)**, 953-970.

21. Zhang, G.P., and Hu, M.Y.: "Neural Network Forecasting of the British Pound/US Dollar Exchange Rate," *Omega*, (1998) **26**(4), 495-506.

22. Prybutok, V.R., Yi, J., and Mitchell, D.: "Comparison of Neural Network Models with ARIMA and Regression Models for Prediction of Houston's Daily Maximum Ozone Concentrations," *European Journal of Operational Research*, (2000) **122**, 31-40.

23. Kolehmainen, M., Martikainen, H., and Ruuskanen, J.: "Neural Networks and Periodic Components Used in Air Quality Forecasting," *Atmospheric Environment*, (2001) **35**, 815-825.

24. Shih, Y.: *Neuralyst User's Guide*, Cheshire Engineering Corporation, Pasadena, CA (1994).

25. Gorr, L.: "Research Prospective on Neural Network Forecasting," *International Journal of Forecasting*, (1994) **10**, 1-4.

26. Nelson, M., Hill, T., Remus, T., and O'Connor, M.: "Time Series Forecasting Using Neural Networks: Should the Data be Deseasonalized First?," *Journal of Forecasting*, (1999) **18**, 359-367.

27. Zhang, G.P. and Qi, M. "Predicting Consumer Retail Sales Using Neural Networks," *Neural Networks in Business: Techniques and Applications,* Smith and Gupta (eds.), Idea Group Publishing, Hershey, PA (2002), 26-40.

28. Granger, C.W. J.: "Strategies for Modeling Nonlinear Time-Series Relationships. *The Economic Record*," (1993) **69**, 233-238.

29. Hoptroff, R.G.: "The Principles and Practice of Time Series Forecasting and Business Modeling Using Neural Networks," *Neural Computing and Applications*, (1993) **1**, 59-66.

30. Ashley, R.: "Statistically Significant Forecasting Improvements: How Much Out-of – Sample Data is Likely Necessary?," *International Journal of Forecasting*, (2003) **19**, 229-239.

VITA

MATTHEW SCOTT BOLEN

PERSONAL DATA

Name	Matthew Scott Bolen
Mailing Address	91 East Slatestone Circle The Woodlands, TX 77382 E-mail: mbolen@tamu.edu

EDUCATION

August 2003 – August 2005	Texas A&M University College Station, TX 77382 M.S. Petroleum Engineering
August 1999 – August 2003	Virginia Polytechnic & State University Blacksburg, VA 24061 B.S. Chemical Engineering

EXPERIENCE

Summer 2004	Halliburton Energy Services Longview, TX